REVIEW OF EDUCATIONAL RESEARCH

Official Publication of the American Educational Research Association
Contents are listed in the Education Index.

Copyright 1963
By National Education Association of the United States, Washington, D.C.

Volume XXXIII, No. 2 **April 1963**

Guidance, Counseling, and Personnel Services

Reviews the literature for the three-year
period since the issuance of Vol. XXX, No. 2, April 1960.

TABLE OF CONTENTS

Page

Chapter

Foreword 142

I. Foundations 143
CARROLL H. MILLER, *Northern Illinois University, DeKalb, Illinois*

II. Elementary and Secondary School Programs 152
HAROLD J. REED, *County of Los Angeles, Office of the Superintendent of Schools, Los Angeles, California*
BUFORD STEFFLRE, *Michigan State University, East Lansing, Michigan*

III. Higher Education Programs 163
ESTHER LLOYD-JONES, *Columbia University, New York, New York*
MARGARET RUTH SMITH, *Wayne State University, Detroit, Michigan*

IV. Selection, Preparation, and Professionalization of Specialists 171
ROBERT O. STRIPLING, *University of Florida, Gainesville, Florida*
JAMES L. LISTER, *University of Florida, Gainesville, Florida*

V. Counseling 179
ROBERT CALLIS, *University of Missouri, Columbia, Missouri*

VI. The Appraisal Function 188
LEO GOLDMAN, *Brooklyn College, Brooklyn, New York*

Chapter *Page*

VII. Vocational Development Process 197

GARRY R. WALZ, *University of Michigan, Ann Arbor, Michigan*

VIII. Group Procedures 205

E. WAYNE WRIGHT, *Utah State University, Logan, Utah*

IX. Program Evaluation 214

C. H. PATTERSON, *University of Illinois, Urbana, Illinois*

Index 225

This issue of the REVIEW was prepared by the
Joint Committee of the American Educational Research
Association and the American Personnel and Guidance Association

BUFORD STEFFLRE, *Chairman*, Michigan State University, East Lansing, Michigan
DOUGLAS D. BLOCKSMA, Consulting Psychologist, Grand Rapids, Michigan
R. STANLEY GEX, Eastern Michigan University, Ypsilanti, Michigan
KENNETH H. PARKER, Lamphere Public Schools, Madison Heights, Michigan
EDWARD C. ROEBER, University of Michigan, Ann Arbor, Michigan
MARGARET RUTH SMITH, Wayne State University, Detroit, Michigan
GARRY R. WALZ, University of Michigan, Ann Arbor, Michigan
FRANK B. WOMER, University of Michigan, Ann Arbor, Michigan

with the assistance of

ROBERT CALLIS, University of Missouri, Columbia, Missouri
LEO GOLDMAN, Brooklyn College, Brooklyn, New York
JAMES L. LISTER, University of Florida, Gainesville, Florida
ESTHER LLOYD-JONES, Columbia University, New York, New York
CARROLL H. MILLER, Northern Illinois University, DeKalb, Illinois
C. H. PATTERSON, University of Illinois, Urbana, Illinois
HAROLD J. REED, County of Los Angeles, Office of the Superintendent of Schools, Los Angeles, California
ROBERT O. STRIPLING, University of Florida, Gainesville, Florida
E. WAYNE WRIGHT, Utah State University, Logan, Utah

FOREWORD

This is the eleventh issue of the REVIEW concerned with the general area of guidance and counseling, and it is the second joint effort of the American Educational Research Association and the American Personnel and Guidance Association. This issue is the first to give systematic recognition to the fact that the school counselor increasingly is seen as a member of a pupil personnel team, who works closely with school psychologist, school nurse, and school social worker. In setting the field of guidance and counseling within the pupil personnel framework, the authors distinguished between the techniques of the related pupil personnel specialists and the rationale for their employment, the administrative framework within which they work, and their preparation. As a consequence, this newer emphasis will be found primarily in Chapter I, "Foundations"; Chapter II, "Elementary and Secondary School Programs"; and Chapter IV, "Selection, Preparation, and Professionalization of Specialists."

With but two exceptions, the topics of this issue are similar to those of the April 1960 issue. The chapter on higher education programs, similar to that in the 1954 issue, makes visible the fact that organizations at this level usually differ from those at the elementary or the secondary school level. The chapter on foundations considers not only the philosophical basis for guidance, counseling, and personnel services but also the sociological and cultural framework.

The guidance and counseling uses of appraisal data are considered in the chapter on the appraisal function; whereas research on the more technical aspects of tests has been evaluated in the February 1962 issue, *Educational and Psychological Testing*. Although the Committee attempted to define the content of the chapters to avoid needless duplication of material, some research articles were sufficiently broad in scope to merit discussion in more than one context. In such cases, the varying emphases seem to justify multiple inclusion.

The Chairman wishes to express appreciation to the members of the Committee and to the additional contributors who wrote the various chapters.

Buford Stefflre, *Chairman,*
Joint Committee on
Guidance, Counseling, and Personnel Services

CHAPTER I

Foundations

CARROLL H. MILLER

For the purpose of this chapter, the term *foundations* will be used to mean the societal, cultural, and philosophical context from which are derived, or to which are related, purposes and practices of student personnel and guidance. Since this is the first time that the term foundations in this broad sense has constituted a chapter topic in an issue of the REVIEW devoted to the area of counseling and guidance, comparisons with findings of earlier issues are not possible. However, philosophical foundations were reviewed by Wilkins and Perlmutter in the April 1960 REVIEW. The need for general principles and theoretical foundations reported by these authors is still evident. There is, however, some indication of a rising interest in existentialism, which will be considered in this chapter.

Society and Culture

College Students

A recent volume prepared for the Society for the Psychological Study of Social Issues (Sanford, 1962) provided critical reviews of research and some new data on such topics as characteristics of student populations, student society and culture, the relation of student performance to educational objectives, and the social context of higher education. Sprague (1960) reported recent research on motivation, values, attitudes, and campus cultures. Characteristics of freshmen classes in 128 colleges and universities that are members of the College Entrance Examination Board were summarized as a series of profiles (College Entrance Examination Board, 1963). A report by Hall and Craigie (1962) on students who borrowed through the National Defense Student Loan program included data on age, marital status, family income, major field of study, and occupational plans. A five-year listing of research projects supported by the Cooperative Research Program (U.S. Department of Health, Education, and Welfare, Office of Education, 1962) reflected limited interest in investigation of college society and culture. For example, of the 416 projects listed, only 12 were studies of college environments or of attitudes and values of college students.

Additional References: Blaesser and Crookston (1960); Goldsen and others (1960); Heist (1960); Mueller (1960); Pace (1960); Riesman (1959); Wilk (1960).

School-Age Youth

The report of the Commission on Guidance in American Schools (Wrenn. 1962) interpreted guidance in the American setting. This important material on social change and changes in the schools not only described the present situation in broad terms but also offered projections and recommendations. Coleman (1962) made the final report of a noteworthy investigation of adolescents in nine high schools, with emphasis on status systems. The insulation of adolescent society from the adult world suggested by the findings poses a problem for which the investigator saw two solutions: (a) bring the adolescent back to the home or (b) accept adolescent society as it is and use it for educational purposes. An extensive investigation was reported by Reiss and Rhodes (1959) involving almost 22,000 adolescents in the Nashville metropolitan area. With its focus upon factors related to conformity to or deviation from cultural goals, the immediately obvious implications relate to delinquency, especially if delinquent behavior is defined as norm-violating behavior, as it was in the National Education Association project (Kvaraceus and Miller, 1959). But the Nashville study also helps to illuminate attitudes of adolescents in general toward their own and the larger American culture.

Havighurst and others (1962), in *Growing Up in River City*, reported a longitudinal and intensive study of one community. Almost 500 pupils who in 1951 were in the sixth grade (chiefly in public schools) were followed through 1960. The study covers many facets: the community itself, abilities, achievement in school, adjustment, delinquency, marriage, the first two years in college or in early employment, and early adult status and mobility. The threads of inquiry are held together by two general questions: (a) How well did the boys and girls succeed in the tasks of growing up? (b) How well did the community provide for their needs? A very different kind of report was made in *High School Students Speak Out* (Mallery, 1962). This book is a statement, largely in the students' own words as gathered by classroom visits and in interviews, of their experiences in high school. The sample included students in eight schools in three regions. Using still a different approach, Friedenberg (1959) considered the adolescent's problem of defining himself in the social conditions of contemporary American life. His analysis, made largely in a Sullivanian framework, led him to believe that adolescence as a developmental process is becoming obsolete.

Additional References: Dugan (1960); Findley (1960).

Values

Most of the studies mentioned thus far at both school and college levels gave some attention to the matter of values, but some others were more specifically so oriented. Klinger (1962) studied the moral values of

U.S. and foreign male students, including Arab Moslems, Chinese non-Christians, Indian Hindus, South American Roman Catholics, and Turkish Moslems. Cross-group value structures as well as those more characteristic of single groups were described. Warnath and Fordyce (1961) found that among five groups of college freshmen who had declared majors, the humanities group separated from others on an aesthetic-values scale, and the business group on a material-values scale. Maehr and Stake (1962) found that students who persisted in a theological seminary scored lower on aesthetic values than those who withdrew prior to graduation. Prince (1960) found significant differences in value scores among high school students in public, private, and religious schools; but within each school differences between groups of freshmen and seniors were not important. Both groups resembled the teachers in the schools in their value patterns. A study of values of students in seven Catholic high schools indicated that attitudes of lower class parents and children were most favorable toward belief in God and toward belief in the dignity of each individual, while attitudes toward these beliefs were most unfavorable among upper class parents and children (Del Grande, 1959).

Additional References: Ginzberg (1961); McArthur (1960); Simpson and Simpson (1960).

The Culturally Deprived

The present strong interest in inequality of opportunity afforded in varied community settings is clearly indicated by the response to reports such as *Slums and Suburbs* (Conant, 1961) and *Social Dynamite* (National Committee for Children and Youth, 1961). Although problems of inequality are related to the topic of foundations, the research content of available reports is often minimal.

A review of 142 projects listed in *Social Dynamite* suggested that while some have been conceived as primarily research undertakings, many were simply reports of ongoing programs. Riessman (1962) said that the Higher Horizons program of New York City has rendered an important service in showing that the culturally deprived can be educated. Nevertheless, he raised the problem of whether the results reflect a Hawthorne effect. However, in the report by Havighurst and others (1962) mentioned earlier, the authors commented that although there were several types of disadvantage, that of low socioeconomic status was foremost. The research difficulties in this area are manifold; even the very term *culturally deprived* is not easily defined. Any attempt at a summary of findings would be inappropriate at this time, especially since many of the projects are still in progress.

Additional References: Impellizzeri (1960); Krugman and Impellizzeri (1960).

Philosophical Trends

One of the most noticeable characteristics of the literature related to foundations is a lack of philosophical criticism. However, a few contributions are well worth noting. The social thinking of five pioneers in the field was interpreted by Rockwell and Rothney (1961). Frank Parsons was presented as a utopian social reformer and Jessie B. Davis as an advocate of the social gospel; Anna Y. Reed and Eli Weaver were identified with social Darwinism; and Davis S. Hill was seen as more concerned than the others with applying scientific research to the study of the individual and to social problems in general.

Noting that traditional philosophy has contributed to guidance certain unanalyzed assumptions and unquestioned beliefs, Gowin (1961) pointed out that attempts to formulate a philosophy of personnel work have often resulted in an eclecticism of realism, idealism, rationalism, and experimentalism. Most counselors, he noted, do not appear to be familiar with existentialism and logical positivism. Existentialists tend to reject the social order, while positivists ignore rather than reject it. Gowin speculated about the kinds of answers these two orientations might give the counselor.

Wise (1962) found in existentialist thinking two concepts of importance to guidance: (a) the concept of responsibility that sees the world as the structure of meaningful relationships in which the person exists and (b) the concept of freedom that derives from the concept of responsibility. Morris (1959) expressed concern about the relation of guidance and freedom. He suggested that we might expect guidance to be on the side of individual self-realization but that, actually, it appears to support the "organization man" principle.

The thread of existentialist thinking which seems increasingly apparent is evident in the latest book by Rogers (1961). Starting with experience in therapy, Rogers explored several philosophical questions. Several examples may suggest his tone: On the matter of goals or purposes of the person, he reported that his observations of persons in therapy who were trying to find a way of life led him to think that there is a general pattern emerging that might best be described in the words of Kierkegaard—"to be that self which one truly is." Another example is the paper in which he discussed the difficulties of reconciling logical positivism with the "subjectively oriented existential thinking which . . . seemed to fit so well with my therapeutic experience." Or, stated differently, it is the question, "Persons or science?"

Additional References: Lowe (1959); Pervin (1960); Shoben (1961); Smith (1959, 1961).

Concepts and Purposes

Principles and assumptions of guidance and personnel practice and the concepts and purposes relevant to them are, to say the least, greatly varied.

There seems to be little discernible movement toward any generally accepted synthesis. Although concepts of counseling and program organization are relevant, they are covered in other chapters.

In the Schools

In a statement by Bowles (1959), guidance was conceived primarily as educational guidance involving appraisal of the individual and educational placement. Cox (1961) noted three new emphases in vocational guidance: (a) an internal rather than an external criterion of success, (b) the work life as a developmental process, and (c) the relation of vocational choice and personality. Weitz (1961) interpreted guidance as behavioral change and suggested a model, which he applied to testing and counseling. Mathewson (1961) suggested four dimensions in guidance practice: (a) classroom-centered–specialized, (b) self-integrative–prescriptive, (c) personal-social, and (d) subjective-objective. In a somewhat similar manner, Glanz (1961) identified four patterns: (a) centralized specialism, (b) decentralized generalism, (c) curricular counseling and guidance (group guidance), and (d) human relations and group work (mental health).

In a historical study of guidance in Ohio since 1900, Hummel (1960) found one of the trends to be an emphasis on self-realization and freedom in choice making. Ginzberg (1960) questioned three assumptions commonly made: (a) guidance should deal with the whole person; (b) the counselor has a responsibility for helping the individual to understand himself; and (c) guidance is a specific field, and the counselor an expert. No evidence is available by which to judge the extent to which schools generally have attempted to clarify and state their purposes, but a study in Ohio found that only 32 percent of 958 schools surveyed indicated that they had formulated statements of guidance philosophy—statements which varied considerably in orientation (Peters, 1961).

Additional References: Barry and Wolf (1962); Council of Chief State School Officers (1960); Cribbin (1961); Hilton (1959).

In Colleges and Universities

The past three years have yielded few critical reconsiderations of purposes and policy of student personnel work. Wise (1961) suggested that for the past 30 years the student personnel "movement" has neglected both historical roots and new developments so that a cult has arisen "isolated from many of the central influences in American education, particularly higher education." Mueller (1961) saw student personnel work as "an applied social science with its roots in psychology and sociology." In one sense, it is a special aspect of teaching, with functions related to the objectives of higher education but differentiating the specific objectives of its own functions, which center about the integration of campus life. In his latest book, Williamson (1961) wrote of student personnel work "as a tangible expression in the program of a basic philosophy." He described five aspects of this point of view: (a) "holism," or concern with all aspects of development; (b) recognition of "the unique individuality of each individual"; (c) belief that in the education of some students, classroom teaching is not enough; (d) use of methods and relationships in an educative rather than an authoritarian way; and (e) utilization of new knowledge from many fields. The chief emphasis of Williamson's book, which was concerned with a philosophy of administration, actually is beyond the scope of this chapter.

Summary

Research in the area of foundations still lacks integration in terms of any large concept or set of concepts. Clearly, one immediate need is a definition that will set limits on the area of foundations. Within the definition arbitrarily adopted in this chapter, one focus of research that seems to be emerging centers about peer society and culture as they exist in schools and colleges. This focus is clearer now at the college than at the secondary school level.

A conspicuous feature of the literature concerned with the college level has been the appearance of broad syntheses of research in certain areas. At the secondary school level, reports of specific problems by individual investigators have been more common. Statements of concepts of guidance in schools and challenges of these have been so varied that they defy an interpretation of trends; they seem to reflect a welter of current opinion and practice.

On the level of philosophic inquiry, eclecticism appears to be less dominant than in the past, and some trend toward an interest in existentialism is appearing. Perhaps this reflects an influence from emphasis on therapeutic counseling, or it may reflect existential elements in the broad context of contemporary culture.

Bibliography

BOWLES, FRANK H. "The Nature of Guidance." *Personnel and Guidance Journal* 38: 112-19; October 1959.

COLEMAN, JAMES S. *The Adolescent Society.* New York: Free Press of Glencoe, 1962. 368 pp.

COLLEGE ENTRANCE EXAMINATION BOARD. *Manual of Freshman Class Profiles.* New York: the Board, 1963. 600 pp.

CONANT, JAMES B. *Slums and Suburbs.* New York: McGraw-Hill Book Co., 1961. 147 pp.

COX, RACHEL DUNAWAY. "New Emphases in Vocational Guidance." *Vocational Guidance Quarterly* 10: 11-15; Autumn 1961.

Del Grande, Sister Mary Vera. *A Study of the Values of Catholic High School Students of Differing Socio-Economic Backgrounds and the Relationship of These Values to Those of Their Parents.* Doctor's thesis. St. Louis: St. Louis University, 1959. 454 pp. Abstract: *Dissertation Abstracts* 20: 3509; No. 9, 1960.

Friedenberg, Edgar Z. *The Vanishing Adolescent.* Boston: Beacon Press, 1959. 144 pp.

Ginzberg, Eli. "Guidance—Limited or Unlimited." *Personnel and Guidance Journal* 38: 707-12; May 1960.

Glanz, Edward C. "Emerging Concepts and Patterns of Guidance in American Education." *Personnel and Guidance Journal* 40: 259-65; November 1961.

Gowin, D. B. "Prospects for Philosophic Inquiry in Guidance." *School Review* 69: 191-205; Summer 1961.

Hall, Robert C., and Craigie, Stanton. *Student Borrowers: Their Needs and Resources.* U.S. Department of Health, Education, and Welfare, Office of Education, Publication OE-55011. Washington, D.C.: Government Printing Office, 1962. 80 pp.

Havighurst, Robert J., and others. *Growing Up in River City.* New York: John Wiley & Sons, 1962. 189 pp.

Hummel, Dean Loil. *Guidance in Ohio: Its Historical Development.* Doctor's thesis. Columbus: Ohio State University, 1960. 343 pp. Abstract: *Dissertation Abstracts* 21: 2984; No. 10, 1961.

Klinger, M. Robert B. "Moral Values Across Cultures." *Personnel and Guidance Journal* 41: 139-43; October 1962.

Kvaraceus, William C., and Miller, Walter B. *Delinquent Behavior.* Washington, D.C.: National Education Association, 1959. Part I, "Culture and the Individual," 147 pp.

Maehr, Martin L., and Stake, Robert E. "The Value Patterns of Men Who Voluntarily Quit Seminary Training." *Personnel and Guidance Journal* 40: 537-40; February 1962.

Mallery, David. *High School Students Speak Out.* New York: Harper & Row, 1962. 165 pp.

Mathewson, Robert H. "School Guidance: A Four-Dimensional Model." *Personnel and Guidance Journal* 39: 645-49; April 1961.

Morris, Van Cleve. "Conformity, Rebellion, and the Authentic Life." *Teachers College Record* 61: 46-50; October 1959.

Mueller, Kate Hevner. *Student Personnel Work in Higher Education.* Boston: Houghton Mifflin Co., 1961. 570 pp.

National Committee for Children and Youth. *Social Dynamite: The Report of the Conference on Unemployed, Out-of-School Youth in Urban Areas.* Washington, D.C.: the Committee, 1961. 265 pp.

Peters, Herman J., project director. *Guidance in Ohio.* Columbus: Ohio State Department of Education, 1961. 230 pp.

Prince, Richard. "Values, Grades, Achievement, and Career Choice of High School Students." *Elementary School Journal* 60: 376-84; April 1960.

Reiss, Albert J., Jr., and Rhodes, Albert Lewis. *A Socio-Psychological Study of Conforming and Deviating Behavior Among Adolescents.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Program, Project No. 507. Iowa City: State University of Iowa, 1959. 656 pp.

Riessman, Frank. *The Culturally Deprived Child.* New York: Harper & Brothers, 1962. 140 pp.

Rockwell, Perry J., Jr., and Rothney, John W. M. "Some Social Ideas of Pioneers in the Guidance Movement." *Personnel and Guidance Journal* 40: 349-54; December 1961.

Rogers, Carl R. *On Becoming a Person.* Boston: Houghton Mifflin Co., 1961. 420 pp.

Sanford, Nevitt, editor. *The American College: A Psychological and Social Interpretation of the Higher Learning.* New York: John Wiley & Sons, 1962. 1084 pp.

Sprague, Hall T., editor. *Research on College Students.* Boulder, Colo.: Western Interstate Commission for Higher Education, 1960. 188 pp. (Offset)

U.S. Department of Health, Education, and Welfare, Office of Education. *Research Projects of the Cooperative Research Program: A Five-Year Summary, July 1, 1956—September 30, 1961.* Publication No. OE-12015. Washington, D.C.: Government Printing Office, 1962. 79 pp.

WARNATH, CHARLES F., and FORDYCE, HUGH R. "Inventoried Values of Entering College Freshmen." *Personnel and Guidance Journal* 40: 277-81; November 1961.

WEITZ, HENRY. "Guidance as Behavior Change." *Personnel and Guidance Journal* 39: 550-62; March 1961.

WILLIAMSON, E. G. *Student Personnel Services in Colleges and Universities.* New York: McGraw-Hill Book Co., 1961. 474 pp.

WISE, W. M. "Student Personnel Work: Future Trends." *Personnel and Guidance Journal* 39: 704-709; May 1961.

WISE, W. M. "Existentialism and Guidance-Personnel Work." *Journal of the National Association of Women Deans and Counselors* 25: 161-65; June 1962.

WRENN, C. GILBERT. *The Counselor in a Changing World.* Washington, D.C.: American Personnel and Guidance Association, 1962. 195 pp.

Additional References

BARRY, RUTH, and WOLF, BEVERLY. *Epitaph for Vocational Guidance.* New York: Bureau of Publications, Teachers College, Columbia University, 1962. 241 pp.

BLAESSER, WILLARD W., and CROOKSTON, BURNS B. "Student Personnel Work—College and University." *Encyclopedia of Educational Research.* Third edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1415-27.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS. *Responsibilities of State Departments of Education for Pupil Personnel Services.* Washington, D.C.: the Council, 1960. 25 pp.

CRIBBIN, J. J. "Essentials and Incidentals in Guidance." *Catholic Educational Review* 59: 608-19; December 1961.

DUGAN, WILLIS E. "The Nature and Variability of High School Graduates—A Basic Factor in School-College Articulation." *North Central Association Quarterly* 35: 181-92; October 1960.

FINDLEY, WARREN G. "Student Personnel Work—Elementary and Secondary." *Encyclopedia of Educational Research.* Third edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1427-33.

GINZBERG, ELI, editor. *Values and Ideals of American Youth.* New York: Columbia University Press, 1961. 338 pp.

GOLDSEN, ROSE K., and OTHERS. *What College Students Think.* Princeton, N.J.: D. Van Nostrand Co., 1960. 240 pp.

HEIST, PAUL. "Diversity in College Student Characteristics." *Journal of Educational Sociology* 33: 279-91; February 1960.

HILTON, M. EUNICE, editor. *New Frontiers in Guidance-Personnel Work.* Syracuse, N.Y.: Syracuse University Press, 1959. 108 pp.

IMPELLIZZERI, IRENE H. "The New York City Talent Preservation Project: An Experiment in Guidance of the Mentally Superior." *Journal of the National Association of Women Deans and Counselors* 23: 99-103; April 1960.

KRUGMAN, MORRIS, and IMPELLIZZERI, IRENE H. "Identification and Guidance of Underachieving Gifted Students." *Exceptional Children* 26: 283-86; February 1960.

LOWE, C. MARSHALL. "Value Orientations—An Ethical Dilemma." *American Psychologist* 14: 687-93; November 1959.

MCARTHUR, CHARLES. "Sub-Culture and Personality During College Years." *Journal of Educational Sociology* 33: 260-68; February 1960.

MUELLER, KATE HEVNER. "The Married Student on the Campus." *College and University* 35: 155-63; Winter 1960.

PACE, C. ROBERT. "Five College Environments." *College Board Review* 41: 24-28; Spring 1960.

PERVIN, LAWRENCE A. "Existentialism, Psychology, and Psychotherapy." *American Psychologist* 15: 305-309; May 1960.

RIESMAN, DAVID. "The Influence of Student Culture and Faculty Values in the American College." *Year Book of Education, 1959: Higher Education.* (Edited by George Z. F. Boreday and Joseph A. Lauwerys.) Yonkers-on-Hudson, N.Y.: World Book Co., 1959. pp. 386-404.

SHOBEN, EDWARD JOSEPH, JR. "Personal Responsibility, Determinism, and the Burden of Understanding." *Personnel and Guidance Journal* 39: 342-48; January 1961.

SIMPSON, RICHARD L., and SIMPSON, IDA HARPER. "Values, Personal Influence, and Occupational Choice." *Social Forces* 39: 116-25; December 1960.

SMITH, M. BREWSTER. "Research Strategies Toward a Conception of Positive Mental Health." *American Psychologist* 14: 673-81; November 1959.

SMITH, M. BREWSTER. "'Mental Health' Reconsidered: A Special Case of the Problem of Values in Psychology." *American Psychologist* 16: 299-306; June 1961.

WILK, ROGER E. "Student Personnel." *Recent Research and Developments and Their Implications for Teacher Education.* Thirteenth Yearbook, American Association of Colleges for Teacher Education. Washington, D.C.: the Association, 1960. pp. 53-63.

CHAPTER II

Elementary and Secondary School Programs

HAROLD J. REED and BUFORD STEFFLRE

IN THE April 1960 REVIEW, the last issue on the organization and administration of guidance services, Dugan noted the paucity of research on guidance policy, staff qualifications, budgets, facilities, and counselor duties. The past three years were no exception.

Some of the areas—facilities, budgets, referrals, and cumulative records—which were reported then as minimal activities have not changed. Emphasis on some elements, such as identification and development of individual potential, elementary guidance services, and, especially, the duties and functions of counselors, is greater now than during the period studied by Dugan. Some areas have appeared in the literature for the first time during the past three years. Data-processing techniques and the increasing reference to *pupil personnel services* instead of *guidance services* are examples.

A greater number of areas have been investigated and reported in recent literature as guidance services have expanded. However, the specific contributions in organization and administration are disproportionately low. Of 380 projects initiated under the Cooperative Research Program (U.S. Department of Health, Education, and Welfare, Office of Education, 1962) during the five-year period from 1956 to 1961, only 22 were classified under counseling and guidance, and only one title suggested a research project in organization and administration of guidance services.

Whether one examines organization and administration practices for validity or as criteria for validating some predictive methodology, one finds that little research is being done. However, the picture may be only a perceptual one. A means-end-means relationship may be operating; and what appears to be a serious research deficiency may, in fact, reflect a healthy state of making administrative decisions on the basis of research findings in techniques, training, personnel selection, and other prerequisite elements in a good educational program. If, on the other hand, it is found from a review of the literature that organization, job functions, and program problems are not being researched and that administrative decisions are being made on an "it-is-agreed" basis, then there is some cause for alarm.

Organization and Administration of Guidance Services

Guidance services as measured by improved counselor-student ratios, by increases in the number of state department supervisors, and by higher counselor standards have been aided by the National Defense Education Act (NDEA) program since 1958 (Crummel, 1961). The number of counselors has increased by more than 50 percent in two years. State department guidance workers have increased by 150 percent.

Current administrative needs may be assessed by noting the use of NDEA funds. The distribution of the 185 district and county NDEA projects reported in California by the California State Department of Education (1962a) included (a) 150 to increase professional staff time, (b) 138 to increase clerical time, and (c) 75 to provide parent conferences.

Another research study reported by the California State Department of Education (1962d) showed the influence of professional credentialing on the quality of counseling. Since 1956, California has required a pupil personnel services credential for those performing counseling services as a major job function. A survey of 81 percent of the schools revealed that 55 percent of all counseling is performed by the 42 percent of counselors who are credentialed. Of those who are not credentialed, 62 percent are taking courses. Students in medium-sized schools are more apt to receive counseling from well-trained and full-time counselors than those in small or large schools.

Types of Counselor Assignments

Some needed research relating to the assignment of counselors has appeared, and much more is needed. When responses of students on counseling wanted and the values of help received were classified by organization of services—that is, full-time counseling, part-time counseling, and counseling by the teacher with no released time—it was found that students in teacher-counseled classes asked for more help than those in the other two organizational systems, but counselees of full-time counselors reported more help received (Thomson, 1960). Loughary (1959) examined some organizational systems and concluded that in large schools where a choice is possible secondary school administrators do not agree on the relative merits of full-time counselors as opposed to teacher-counselors. The rationale for this conclusion, according to Loughary, is that specialists might lose sight of major educational objectives. Advocates of the full-time counselor system argue that specialized skills are necessary.

Without attempting to compare the relative merits of full-time *versus* part-time counseling, the California State Department of Education (1962c) found that in 645 high schools surveyed, 21 percent of the counselors were scheduled full time and 59 percent, half time or more.

Use of Ratios

Ratio is the most frequent technique used for assessing counseling services, but an erroneous assumption is made if ratio is not computed on the basis of credentialed or qualified counselors. The California State Department of Education (1962b) discovered that the ratio of 1 full-time counselor to 412 secondary students in 81 percent of the state's high schools changed to 1 to 745 when only credentialed or qualified counselors were used to compute ratios. Districts under 500 and over 15,000 had a ratio of 1 qualified counselor to 1,300 students, while other districts averaged 1 to 600. Ratios can also be misleading if, for example, one district assigns attendance services to an attendance worker, while another assigns attendance to the counselor. Discipline may be handled by the counselor in one school and by the assistant principal in another. Obviously, comparative ratios have little or no meaning in these and other similar situations.

The Wrenn (1962) report attempted to consider misleading ratio factors by stating that a counseling program exists only when a sufficient number of qualified counselors is present at a school to take care of the counseling needs of all students. According to the American School Counselor Association (ASCA) recommendation, if the counseling program includes counseling students, consulting with teachers, studying characteristics of student population, and coordinating services in the school and with the community, then the ratio in high school should be 1 to 300 and, in elementary school, 1 to 450.

Use of Data-Processing Equipment

A significant contribution to the administration of guidance services, including the utilization of professional personnel, has been the rapid increase in the use of data-processing equipment. Gathering, recording, and using essential pupil information has long been considered a high-priority guidance problem. District programs are demonstrating the effectiveness of machine processing at an accelerating rate (Kinling, 1960). A pilot data-processing center has been in operation for two years in California, following a study that suggested that processing data was more efficient in 10 districts using machines than in 10 districts using manual methods (Grossman, 1962). Machines established their superiority in processing pupil personnel data and utilizing counselor time in clerical duties—29 percent for data processing and 39 percent for schools using manual procedures. It was tentatively concluded that financial economy has been overestimated and that human economy was the real gain.

Special Services

Dukelow (1961) and Hypps (1959) surveyed large districts and found that they were providing psychiatric, psychological, mental health, and special education services and classes for the exceptional child. However, the respondents to Dukelow's survey indicated that 7.5 percent of the pupils could benefit from mental health services and less than 2 percent were receiving help. Jampolsky (1961), as a psychiatric consultant, described a successful elementary school program for disturbed children involving a special class for pupils, group therapy for parents, and a child study program for teachers.

Timing in Introducing Guidance Services

Another problem that requires the assistance of related professional contributions is determining the appropriate time to introduce certain guidance services. Educational and career guidance programs, for example, have been introduced during recent years at grade levels earlier than the twelfth grade, in which they had been traditional (Sinick and Hoppock, 1960). Two elements have been partially responsible: (a) the increasing emphasis on identification of the talented at a level where there is more talent in proportion to size of age group in the total population and (b) the growing acceptance of counseling and psychological services as different from those performed by the teacher.

There is some evidence that research in developmental psychology is influencing program decisions, but not sufficiently. Progress reports on the vocational maturity study of ninth grade boys have demonstrated the value of submitting theories and hypotheses to experimentation (Super, 1960; Super and Overstreet, 1960). The authors concluded that ninth grade boys were ready to explore, but not to make decisions—a finding that suggests the objectives and content of guidance courses. Gribbons (1960) found that youth at intermediate and junior high school age can benefit from information and opportunities to synthesize their understanding and application of concepts of abilities, values, and interests to their own choices.

The trend toward introducing educational and career guidance programs into the elementary and junior high schools has provided another reason to continue discussing the relative contributions of the teacher and specialist to the guidance program. Shimberg (1959) found that when teachers were provided the Educational Testing Service manual and guide, *You: Today and Tomorrow*, 108 eighth grade students learned essential elements in career planning.

Teacher Versus Specialist Participation

Teacher attitudes toward guidance and participation in guidance services are positively and significantly related, according to Stewart (1961). Randomly selected teachers in the state of Washington with high and low attitudes toward guidance had significantly different scores on the *Minnesota Teacher Attitude Inventory*.

Some authorities have contended that the specialist has a function to perform that is uniquely different from that of the teacher. Hatch and

Costar (1961) pointed out that guidance and good teaching are not synonymous. Prevention programs require the services of the specialist, according to Eckerson and Smith (1962a, c). Cottingham and Hopke (1961) surveyed junior high schools and noted that teachers are responsible for guidance functions in core classes and in other group guidance programs, but they felt that the desirable trend is toward the utilization of specialized guidance personnel.

Until recently, most guidance services at the elementary level were performed with a clinical emphasis by school psychologists. However, Wrenn (1962) reported the results of a survey of elementary school counselors, indicating that during the next 20 years 45 percent expected their role to be one of coordinating many counseling duties and helping parents and teachers, while only 23 percent predicted a clinical emphasis, and 32 percent, a combination of the two. Wrenn also reported a similar trend in a Bakersfield, California, study. After a two-year counseling program, the teacher's perception of the counselor's role in psychological services decreased from 70 to 20 percent.

Roles and Functions of Pupil Personnel Workers

The division of labor so characteristic of American industry also characterizes education. A variety of noninstructional services such as the following have been grouped together under pupil personnel services, and specialists have been assigned to perform their respective functions: attendance, guidance, health, psychological, and social work (Fusco, 1961; Council of Chief State School Officers, 1960).

Health Services

Medical health services are increasingly considered part of the pupil personnel team, but not to the same extent as psychological, counseling, and welfare services. The need for health services is critical. Davis (1960) discussed the implications for the administration of school health services of a national survey of 36,000 households. The survey revealed (a) the extent of various physical limitations in children under age 15, (b) the average number (3.5) of acute episodes or illnesses suffered by this age group each year, and (c) the average number of absences from school by sex and place of residence.

Neilson (1960) reported a study of school health services in 250 urban communities and found that medical doctors became involved administratively in proportion to the size of districts—actively in large districts and as consultants in smaller districts. Only 31 percent of the schools were found to have regular guidance-health services.

The school nurse is apparently the responsible health worker (Klein, 1959). Klein found that 95 percent of pupils with health problems were referred to the nurse as were a large number of pupils with the following problems: 77 percent with home problems, 75 percent with mental health problems, 57 percent with social problems, 36 percent with moral and spiritual problems, and 17 percent with vocational problems.

Psychological Services

Growth in psychological services during the last 11 years as measured by membership in the Division of School Psychologists of the American Psychological Association has been a remarkable 800 percent. Higgins (1959), in a rather utopian proposal for a hypothetical medium-sized district, suggested a psychological service staff with a ratio of 1 psychologist for every 900 pupils. Trachtman (1961) pointed out that Chicago, which now has 90 such specialists, would need 450; that New York City with 100, would need 1,000; and that if every qualified psychologist in the country went to work for the schools, the proposed ratio would still not be met.

Self-Perceptions of Counselors

The self-perceptions of counselors may be too psychologically oriented, if the inference from the ASCA study reported by Wrenn (1962) is valid. It was found that all of the 242 secondary counselors and 71 percent of 138 elementary counselors who were ASCA members listed counseling students as a basic activity; and between 30 and 45 percent listed conferences with parents and teachers, test administration, and group guidance. However, only 8 percent listed involvement in curriculum development as a basic activity; and 12 percent, collection and dissemination of vocational information. The same sample of secondary counselors made recommendations for future counselor preparation. Fifty-four percent selected psychology; 23 percent, the next subject most frequently mentioned; and 11 percent, vocational information. One interpretation of these data would suggest that at least this sample of counselors is more psychologically oriented than, for instance, vocationally oriented. This is significant since vocational orientation is frequently listed as one of the most important counselor functions.

Nomenclature of Guidance Services

The nomenclature of guidance services reflects the state of conflict in the profession. A *guidance worker* may be called a *counselor* in one setting, and a *counselor* may be called a *psychologist* in another. Eckerson and Smith (1962a, b, c) discovered that an elementary *guidance consultant* may be variously titled *counselor, visiting teacher, adjustment teacher*, or *counseling teacher*.

Role and Function of the Counselor

The *school counselor* as a title and as a functionary is the chief pupil personnel worker. If the size and policy of the district does not permit the full complement of pupil personnel workers, the counselor is assigned the functions of the other team members.

Two approaches have been taken to a study of the counselor's role and function: (a) The most common approach is the one in which the opinions of those directly and indirectly related to counseling are surveyed, and the results are summarized. These groups include the counselors themselves, students, teachers, parents, counselor trainers, and others. (b) Another approach has been to relate staff personnel to educational objectives, to the needs of children and youth, and to the unique characteristics of the community. Program decisions are then made, followed by an assignment of duties to each member of the staff according to skills and interests. Although the latter approach is not a common one, judging by the few available references, Sorenson (1961) related counselor functions to three school objectives, with the intellectual one given highest priority. Of the other two—socialization and development of the individual—the latter is more important, harder to achieve, and more likely to be neglected. The counselor is committed to the individual in a way that no one else is. He gives information, but refrains from determining how it should be used. He appraises the student, but does not judge him. He works with disciplinary problems, but does not punish; helps plan, but does not make choices.

The approach to a study of counselor role and function represented by a survey of practices is currently the more popular one, as evidenced by the greater number of references. The four major functions performed by the school counselor according to Wrenn (1962) are (a) counsel students; (b) consult with teachers, parents, and administrators; (c) study and interpret changing facts about student population; and (d) coordinate counseling services. Malcolm (1961) contended that clarification of roles and functions in such areas as counseling, consulting, and research will result in more precise terminology. The California State Department of Education (1962b) found that counselors divided their time as follows: counseling students, 39 percent; noncounseling duties, 38 percent; consulting with parents, 11 percent; consulting with teachers, 10 percent; and liaison with the community, 2 percent.

The ASCA study reported by Wrenn (1962) revealed that of the full-time counselors, 35 percent spent one-half of their time or more counseling students; 30 percent spent two periods or more consulting with teachers and parents; and 70 percent spent one or two periods doing group work with students.

One function that to date appears to be unacceptable to counselors is handling discipline, a function which one-half of the administrators sampled by Cortale (1961) felt that they should perform; the administrators also felt that one-third of the teachers would accept the counselors more if they did handle discipline. Opstad (1960) apparently recommended that counselors accept discipline referrals, but with certain restrictions such as not acting in a punitive, threatening, or investigating role. The Bakersfield, California, study reported by Wrenn (1962) found that teachers did not accept the counselors' heavy emphasis on *clinical* counseling and that they expected counselors to assume a role in discipline.

The opinions of students were surveyed by Heilfron (1960) and by Gibson (1962). Heilfron found that students perceived counselors as working primarily with abnormal cases. The areas of deficiency identified by Gibson could be interpreted as functions according to students' perceptions. Bergstein and Grant (1961) surveyed 200 parents distributed over four grade levels and found that they perceived the counselor favorably, more in educational and vocational areas than in personal-social problems.

A significant problem in guidance administration is the degree of agreement on the counselor's role between the counselor and his administrators. A Q-sort by Schmidt (1962) showed essential agreement, but counselors and administrators may have been satisfied about different things. Chenault and Seegars (1962) compared results on Leary's *Interpersonal Check List* and concluded that the significant factor in differences between administrators and counselors may not have been expectations of each other's role but expectations of each other as persons.

Two current studies will have a great impact on role and function definition. The standards and role studies of the Association for Counselor Education and Supervision and American School Counselor Association (1962) will attempt to define counselor role and function by surveying the opinions and practices of counselors and counselor trainers. The assumption seems to be that those performing the counseling service should set standards and develop criteria. Those who would define the counselor's role by determining the role that guidance or pupil personnel services should play in the total educational program would not necessarily agree that the role of the counselor is what the counselor does. Practically all authorities and counselor surveys, for example, list vocational placement as a counselor function (Wrenn, 1962), but the number of placement services indicates that the service has not become policy in most schools.

Issues Relating to Coordination of Pupil Personnel Services

The trend toward organizing pupil personnel workers as a team will apparently continue but not without problems. Attempts to view the total pupil personnel program in an integrative fashion were made by Johnson, Stefflre, and Edelfelt (1961) and by Rosecrance and Hayden (1960). Both texts synthesized previous research on such matters as incidence of specialists, their duties, training, and suggested ratios.

Roles and functions of the various team members show much duplication when definitions by each group are compared. Job titles and job functions become confused in research, discussions, and association activities. If counselors counsel, what does a social worker do when working with an individual?

The Commission on Guidance in American Schools (Wrenn, 1962) recommended elimination of the term *guidance*, stating that it represents a point of view. McCully (1962) also proposed that the term *counseling services* replace the term *guidance services*. Hoyt (1962), however, disagreed. He proposed a constellation of guidance services, but only in a school setting; whereas McCully would accept a team of workers representing pupil personnel services, but would feel that it would be futile to try to bring them together in a profession. In fact, counselors could not form a profession until the following criteria are met: counseling differentiated from other social service occupations, standards accepted for selection and training, means developed for accrediting training institutions, certification standards established, counseling services improved, and a code of ethics enforced.

A critical problem largely unresearched and otherwise briefly treated is the position of pupil personnel services in the organization and management structure of the school. Knapp and Denny (1961) expressed some concern for the staff and consultant nature of guidance workers and about the tendency for counselors to work with principals rather than teachers, which has not been evaluated. Fusco (1961) observed from surveying eight unified districts that great variations existed in services performed, organizational patterns, ratio, and job titles. The director of pupil personnel services was found, in general, to have line responsibility for only health consultants and other specialists. Nurses and counselors were line-responsible to the principals; and, in one district, the nurses and dental hygienists were line-responsible to both the principal and the staff specialist.

Summary

In summary, it can be said that school guidance programs have continued to grow and to mature. However, most authorities seem to imply that it is necessary to define more clearly the roles of each member of the guidance or pupil personnel team or to combine them into one service under one discipline. Little progress appears to have been made at this point. It will not be made until all pupil personnel subdisciplines supplement their own efforts at role definition by working together toward a common end and as a total group join with other educational departments and administrators in determining the unique contribution of pupil personnel services to the total educational program.

Bibliography

Association for Counselor Education and Supervision and American School Counselor Association. *Counselor Education: A Progress Report on Standards.* Washington, D.C.: American Personnel and Guidance Association, 1962. 90 pp.

Bergstein, Harry B., and Grant, Claude W. "How Parents Perceive the Counselor's Role." *Personnel and Guidance Journal* 39: 698-703; May 1961.

California State Department of Education, Bureau of Pupil Personnel Services. *Annual Narrative Report, Title V-A of the National Defense Education Act of 1958.* Sacramento: the Department, 1962. 35 pp. (a)

California State Department of Education, Bureau of Pupil Personnel Services. *Counselor-Student Ratios for California High School Counselors.* Research Brief No. 3, 1962. Sacramento: the Department, 1962. 6 pp. (b)

California State Department of Education, Bureau of Pupil Personnel Services. *The Organization of Counseling Services in California High Schools.* Research Brief No. 1, 1962. Sacramento: the Department, 1962. 8 pp. (c)

California State Department of Education, Bureau of Pupil Personnel Services. *Professional Preparation of California High School Counselors.* Research Brief No. 2, 1962. Sacramento: the Department, 1962. 8 pp. (d)

Chenault, Joann, and Seegars, James E., Jr. "The Interpersonal Diagnosis of Principals and Counselors." *Personnel and Guidance Journal* 41: 118-22; October 1962.

Cortale, Michael J. "Counselors and Discipline." *Personnel and Guidance Journal* 39: 349-51; January 1961.

Cottingham, Harold F., and Hopke, William E. *Guidance in the Junior High School.* Bloomington, Ill.: McKnight & McKnight Publishing Co., 1961. Chapter 5, "Organizing the Junior High School for Guidance," pp. 62-92.

Council of Chief State School Officers. *Responsibilities of State Departments of Education for Pupil Personnel Services.* Washington, D.C.: the Council, 1960. 25 pp.

Crummel, Robert A. "Analysis of Standards for Guidance Programs Under NDEA." *School Life* 44: 22-24; November-December 1961.

Davis, Roy L. "Quality in School Health Administration." *National Elementary Principal* 39: 6-9; February 1960.

Dukelow, Donald A. "Mental Health Services in Large City Schools: An Inquiry Concerning Mental Health Personnel Serving City Public School Systems." *Journal of School Health* 31: 75-82; March 1961.

Eckerson, Louise Omwake, and Smith, Hyrum M. "Elementary School Guidance: The Consultant." *School Life* 44: 27-31; July 1962. (a)

Eckerson, Louise Omwake, and Smith, Hyrum M. "Guidance in Elementary Schools: Successful Practices." *School Life* 44: 16-19; June 1962. (b)

Eckerson, Louise Omwake, and Smith, Hyrum M. "Guidance in the Elementary School." *School Life* 44: 13-16; May 1962. (c)

Fusco, Gene C. *Organization and Administration of Pupil Personnel Service Programs in Selected School Systems.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1961, No. 22. Washington, D.C.: Government Printing Office, 1961. 67 pp.

Gibson, Robert L. "Pupil Opinions of High School Guidance Programs." *Personnel and Guidance Journal* 40: 453-57; January 1962.

Gribbons, Warren D. "Evaluation of an Eighth Grade Group Guidance Program." *Personnel and Guidance Journal* 38: 740-45; May 1960.

Grossman, Alvin. "A Report of a Study—Processing Pupil Personnel Data." *Bulletin of the California State Department of Education* 31: 1-116; March 1962.

HATCH, RAYMOND N., and COSTAR, JAMES W. *Guidance Services in the Elementary School.* Dubuque, Iowa: Wm. C. Brown Co., 1961. 194 pp.

HEILFRON, MARILYN. "The Functioning of Counseling as Perceived by High School Students." *Personnel and Guidance Journal* 39: 133-36; October 1960.

HIGGINS, CONWELL. "Prospectus: Psychological Service Staff." *School Executive* 78: 42-43; July 1959.

HOYT, KENNETH B. "Guidance: A Constellation of Services." *Personnel and Guidance Journal* 40: 690-97; April 1962.

HYPPS, IRENE C. "The Role of the School in Juvenile Delinquency Prevention with Especial Reference to Pupil Personnel Services." *Journal of Negro Education* 28: 318-28; Summer 1959.

JAMPOLSKY, GERALD G. "How the Public Schools Can Assist in the Promotion of Mental Health." *School Counselor* 8: 111-18; May 1961.

JOHNSON, WALTER F.; STEFFLRE, BUFORD; and EDELFELT, ROY A. *Pupil Personnel and Guidance Services.* New York: McGraw-Hill Book Co., 1961. 407 pp.

KINLING, WILLIAM J. "Dissemination of Guidance Information Using Data-Processing Equipment." *Personnel and Guidance Journal* 39: 220-21; November 1960.

KLEIN, RUTH A. "School Nurse as a Guidance Functionary." *Personnel and Guidance Journal* 38: 318-21; December 1959.

KNAPP, DALE L., and DENNY, EARL W. "The Counselor's Responsibility in Role Definition." *Personnel and Guidance Journal* 40: 48-50; September 1961.

LOUGHARY, JOHN W. "Some Considerations Regarding Full-Time Counselor Versus Teacher-Counselor Assignments." *Educational Administration and Supervision* 45: 199-205; July 1959.

McCULLY, C. HAROLD. "The School Counselor: Strategy for Professionalization." *Personnel and Guidance Journal* 40: 681-89; April 1962.

MALCOLM, DAVID DONALD. "The Image of the Future Secondary School Guidance Program." *Journal of Secondary Education* 36: 72-78; February 1961.

NEILSON, ELIZABETH A. "Analytical Study of School Health Service Practices in the United States." *Journal of School Health* 30: 353-59; November 1960.

OPSTAD, PAUL E. "The Role of the Counselor in Discipline." *Education* 81: 210-12; December 1960.

ROSECRANCE, FRANCIS C., and HAYDEN, VELMA D. *School Guidance and Personnel Services.* Boston: Allyn & Bacon, 1960. 373 pp.

SCHMIDT, LYLE D. "Concepts of the Role of Secondary School Counselors." *Personnel and Guidance Journal* 40: 600-605; March 1962.

SHIMBERG, BENJAMIN. "Classroom Guidance: Hope or Hoax?" *Journal of the National Association of Women Deans and Counselors* 23: 3-10; October 1959.

SINICK, DANIEL, and HOPPOCK, ROBERT. "Research by States on the Teaching of Occupations." *Personnel and Guidance Journal* 39: 218-19; November 1960.

SORENSON, ALMA GARTH. "On the Functions of a School Counselor." *Journal of Secondary Education* 36: 89-91; February 1961.

STEWART, JAMES A. "Factors Influencing Teacher Attitudes Toward and Participation in Guidance Services." *Personnel and Guidance Journal* 39: 729-34; May 1961.

SUPER, DONALD E. "The Critical Ninth Grade: Vocational Choice or Vocational Exploration." *Personnel and Guidance Journal* 39: 106-109; October 1960.

SUPER, DONALD E., and OVERSTREET, PHOEBE L. *The Vocational Maturity of Ninth Grade Boys.* New York: Bureau of Publications, Teachers College, Columbia University, 1960. 212 pp.

THOMSON, MURIEL. "Students' Evaluations of Guidance Services." *California Journal of Educational Research* 11: 167-69; September 1960.

TRACHTMAN, GILBERT M. "New Directions for School Psychology." *Exceptional Children* 28: 159-63; November 1961.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE, OFFICE OF EDUCATION. *Research Projects of the Cooperative Research Program: A Five-Year Summary, July 1, 1956—September 30, 1961.* Publication No. OE-12015. Washington, D.C.: Government Printing Office, 1962. 79 pp.

WRENN, C. GILBERT. *The Counselor in a Changing World.* Washington, D.C.: American Personnel and Guidance Association, 1962. 195 pp.

CHAPTER III

Higher Education Programs

ESTHER LLOYD-JONES and MARGARET RUTH SMITH

ONCE AGAIN it must be reported that research on student personnel programs in higher education in the past three years has been characterized by paucity. The little that has been done will be reported under the following headings: four-year colleges, extended education, and junior colleges. What might be described as sage literature—thoughtful writing by those of broad and long experience—will be briefly summarized.

In the October 1960 issue of the REVIEW, Feder and Siller concluded that there is little evidence concerning the validity of certain stereotypes of organization and function that have been adopted in student personnel practices. They recommended more research involving use of controlled experimentation as well as study of similar institutions with some variation of program in order that the assumptions under which present programs are operating might be validated. Although the tendency to lean heavily on surveys of practices may reveal normative practice and, in effect, establish least common denominators, it may stifle organizational and operational experimentation.

Although the guidance movement is entering into its second half-century, Rothney and Farwell reported in the April 1960 REVIEW that there is general recognition of the need for evaluation of guidance services. However, there is little evidence that this need is being met. Guidance services, like many other services in education, are still offered largely on the basis of hope and faith.

After reviewing research activities pertaining to students and student personnel programs being conducted by 13 agencies and centers for the study of higher education, Robinson and Brown (1961) reported 77 studies in progress or recently concluded. All of these have to do with such matters as student characteristics, attendance, retention, and withdrawal, as well as with aspects of student personnel programs. No studies were reported that looked at student personnel programs as a whole.

Four-Year Colleges

Five studies of student personnel programs deserve mention here. Two doctoral students have attempted to evaluate a single institution's student personnel services. A third study surveyed a group of liberal arts colleges;

another, a group of small colleges. The fifth tackled the problem of studying student personnel work in the context of a college that is professedly Christian.

Beckers (1960) attempted to evaluate the student personnel program on the undergraduate level at the Auburn University School of Education. He took five basic steps to achieve this purpose: (a) The role of student personnel programs in teacher education was defined as a background from which appropriate criteria were derived for appraising student personnel policies and practices. (b) The major elements of the student personnel program in the School of Education were described and analyzed. (c) Student and faculty reactions to the student personnel services were obtained. (d) The student personnel program was evaluated in terms of the criteria derived in the initial step. (e) Recommendations were made on the basis of the findings and conclusions of the study. Data for the investigation were obtained from observations of the student personnel program, from interviews with student personnel staff members, from questionnaires completed by student and faculty members, and from pertinent faculty reports.

Beckers concluded the following: (a) The faculty respondents, in theory, accepted in all or most respects a point of view consistent with democratic values for the institution as a whole. This point of view, however, had not carried over completely to the student personnel program. (b) There was lack of agreement among the faculty respondents about what the over-all purpose of the student personnel program was and what its relationship should be to the teacher preparation program. (c) There was general agreement among the majority of the faculty respondents with respect to the objectives of the individual student personnel service areas. (d) The student personnel program was well supported by administrative leadership, but was lacking in over-all coordination. (e) The availability of student personnel services had been inadequately communicated both to the students as a group and to some faculty members from two departments. (f) The student personnel services generally were inadequate in meeting students' needs. (g) The faculty group that generally had a more favorable opinion of the student personnel program was better acquainted with the student personnel services than was the student group.

The student personnel program at Clark College was studied by Brantley (1960), who attempted not only to investigate intensively Clark's student personnel program, the objectives of its educational program, and the social environments, vocational interests, and problems of Clark College students, but also to make recommendations for the development of the student personnel program.

Brantley assumed that in those instances in which the programs were understood, historically and contemporarily, and in which the environments, interests, and problems of the students were known, an adequate student personnel program could be planned; bases for future research and evaluation could be provided; and information useful to similar institutions could be made available.

Two of his recommendations were (a) that a student personnel worker should understand before his employment any differences that exist between the philosophy of the college and his own educational philosophy and (b) that the administration of the college should clarify for the faculty and the staff the extent to which each personnel service is essential to the total program of the institution.

Inquiry forms were sent by Arbuckle and Kauffman (1959) to 228 "traditional four-year liberal arts colleges." Of the total number, 186, or 84 percent, replied. Six areas were studied: (a) orientation, (b) housing services, (c) health services, (d) financial aid programs, (e) counseling (academic, vocational, and personal), and (f) self-government activities. The evidence from this survey indicates that the modern liberal arts college is by no means unaware of the role of student personnel services in higher education. All the colleges offer some form of most of these services. The colleges reported that although they give most attention to housing and health, they also are quite heavily committed to vocational services. Arbuckle and Kauffman concluded that (a) these colleges are not so antivocational as is sometimes assumed, and (b) counseling is the service which is least effectively provided and concerning which a need for improvement is most expressed.

In a sampling of 16 small liberal arts colleges in the North Central Association area, Scott (1960) compared personnel services, principles, and practices. He found that although the principles of student personnel work were comparable in small and large institutions, the method of implementing the principles varied widely. He suggested that small colleges think seriously about the methods that they use and about the ways in which the services are coordinated with each other and with other areas of the institutions. It should not be assumed that the system of services for students will automatically be well coordinated just because a college is small.

A study by Riggs (1962) of personnel work in the context of a Christian college assumed that student personnel work—like work taught in classrooms—must be intimately related to clearly defined institutional goals. From the statement of institutional goals of De Pauw University, the personnel staff, in weekly meetings and with the help of nationally distinguished thinkers, developed a specific set of goals for the University student personnel program. The basic needs of the kinds of students drawn to and accepted by such a university were carefully studied; and counseling, discipline, perspective, and administration were viewed against this background. This study is cited, not because its research design is clear and its methods rigorous, but rather because its approach moves in the direction advocated by McConnell (1960), Mueller (1961), Brunson (1959), and Williamson (1961).

Extended Education

Simonaitus (1960) undertook a study to determine practices and desirable provisions of the student personnel programs in evening colleges and to examine their influences on enrollment. His findings were based on data from 79 colleges who returned usable questionnaires from among the 120 colleges to whom they were sent. He concluded that most of the provisions considered essential in a personnel program for day college students were also considered essential provisions in a personnel program for evening college students.

Motto (1961) reported from the Flint Extension Center of the University of Michigan a pilot program initiated in 1958 that was oriented toward analyzing student needs, defining the counselor's role, and integrating advisement practices with those of administration. This program asked the following questions: (a) What needs exist for advisement in a graduate study center whose total yearly population exceeds 2,000? (b) What perceptions of advisement are held by graduate students? (c) What student needs, if any, could be met most effectively by a resident advisement program? (d) What range of problems would be presented? (e) What competencies would be required of an adviser who was to service this population? The answers to these questions were sought as a basis for developing a rationale for advisement services in an extension and graduate study center. The results showed the following advisement service requests: (a) educational, 66 percent; (b) vocational, 18 percent; (c) therapeutic, 9 percent; and (d) personal, 7 percent. From these results, Motto drew the conclusion that since one-third of those requesting consideration wanted services other than educational counseling, a counselor to be employed might well have not only a doctorate in clinical or counseling psychology but also such other prerequisites as a practicum training in counseling, public school experience in education, and training in a medical setting.

Junior Colleges

Not too surprisingly, there has been a spate of studies on the junior college level, some of which offer fresh insights for research on other levels.

For the Commission on Student Personnel of the American Association of Junior Colleges, McDaniel (1962) attempted to identify essential student personnel practices in junior colleges. The following salient points were set forth: (a) The purposes of the college are principal determiners of other forces. (b) A student personnel program is an organized pattern of services and activities that makes planned use of incentives and of restrictive controls in order to assist each student to make full use of his college education in progressing toward the goals of the college. (c) Essential information about students must be determined, secured, analyzed, and used. (d) Systematic evaluation of tests used, of grading practices, of class admission standards, and of all college personnel policies is the only protection against mistakes.

Medsker (1960) found in a survey of 76 junior colleges in California that administration and boards of control had not put student personnel services on a par with instruction and that all too often insufficient budget appropriation had been made to insure a full range of such services. The following elements of strength were noted: (a) Administrators recognize the responsibility of a two-year college for an effective student personnel program. (b) Some type and level of counseling program exists in each college. (c) Most colleges have a well-developed student activity program. (d) A good system of academic records is maintained in all colleges. The following elements of weakness were cited: (a) Many institutions lack policy formation, planning, and professional direction of the program. (b) The counseling program in many institutions is inadequate. (c) Little research is conducted to enable the colleges to obtain facts about their students. (d) Two-year colleges make only a limited effort to evaluate their personnel programs.

Glanz, Hayes, and Penney (1959) described the experimental use of a freshman psychology course as furnishing what they believed to be a successful basis of a student personnel program in Boston University Junior College. Although their report made no mention of objective evaluation, Gustad (1959), who disagreed with the experimenters, maintained that the academic program should emphasize one set of goals and that the personnel program should accentuate another although obviously related set. On the other hand, Williamson (1960) agreed with Glanz in stressing that personnel workers should be concerned more, as are the faculty, with the process of education and learning than with the activities underlying administration of services.

Starr's (1960) survey of guidance practices in selected junior colleges in the Northwest revealed that practices stem from four basic concepts: (a) The junior colleges are by nature and, in some instances, by law, preprofessional, vocational, special-education, and general-education institutions dealing with students with varying degrees of maturity, academic development, ambition, and attitudes. (b) Each junior college must be free to develop its program to meet its particular needs. (c) Counseling is important in all phases of the student's life. (d) Counseling and guidance are essential in the junior college.

The following characteristics of programs reported by Starr were noted: (a) Personal-social guidance has not been extensively developed. (b) Lack of available trained counselors is a major weakness. (c) Academic-vocational guidance receives most attention. (d) Extensive use of local community agencies is evident. (e) Extensive use of faculty members often untrained in guidance seems to be customary. (f) Testing programs are offered more consistently than are other services. (g) Little, if any, research or evaluation is evident.

A comparative study of guidance programs in private and public junior colleges was made by Rice (1959). Fifteen private New England and 15 public California junior colleges were studied by means of (a) review and analysis of publications; (b) questionnaires sent to presidents, deans, counselors, and teacher-counselors; (c) visits to campuses for intensive investigations; and (d) personal interviews with officials. He found that guidance aims differed in terms of quality and in relative degree of acceptance. Private junior colleges emphasized the aims of mental health and personal guidance, while public junior colleges stressed the aims of vocational and general guidance. Guidance personnel also differed in many professional respects. For example, public junior college personnel were found to be chosen on the basis of standardized and professionally valid criteria. Compared with private junior college personnel, they had greater experience and training and enjoyed higher professional status and compensatory benefits. Private junior colleges kept more extensive student files of superior over-all quality. Although both private and public colleges offered standard guidance services, neither seemed to have psychiatric help available. No mention was made of the relationship of guidance services to institutional objectives, although both public and private colleges emphasized academic guidance.

Klitzke (1960) pointed to the need for research at the junior college level and indicated important questions dealing with individuals, groups, students and parents, staff, and lay public that need to be answered. He emphasized that student personnel workers need both more research opportunities and more research capabilities.

Issues Underlying Research Needs

Student personnel work cannot contribute effectively unless the overriding goals and value patterns of the institutions are clearly defined and attainable (McConnell, 1960). Educational goals, however, are usually stated in such general form that they do not provide a unifying rationale as a concrete and dependable guide for the institution's activities.

Brunson (1959) made a thoughtful study of the concept of organic administration and of the need for integrating guidance with the rest of the college educational program. She examined by means of case studies the processes used by a variety of liberal arts colleges that were attempting to achieve a higher degree of integration of the separate parts of their education programs.

The structure of the program cannot be divorced with impunity from policy (Williamson, 1961). However, the policy that does and should determine the nature of an enterprise is often so all-pervasive and longstanding that it is frequently taken for granted and sometimes not even understood in detail. To a great extent, education that has been content to organize its programs has failed to study systematically the outcome or products in order to determine the improvements needed.

The personnel worker may first enumerate his goals and projects and then attempt to assess his procedures and accomplishments. If he cannot know the implications of the outcome of his work or evaluate them against the larger world of knowledge, he at least owes to himself the testing of his hypotheses in order that his students and his school may aspire to attain the goals set forth in the program (Mueller, 1961).

The effective administrator expects continuous reorganization and change as improvements are indicated by new and relevant research or evaluation of current programs. The empirical and evaluative approach to organizations and structures would seem to be necessary to avoid the danger of maintaining an inappropriate and outmoded program (Williamson, 1961).

What are the changes taking place about which personnel workers need to be aware when they re-examine and subsequently modify their goals and activities? Wise (1961) has summarized them as follows: (a) extension of educational opportunities to increasing numbers of the population, (b) a rise in the rate of marriage on campus, (c) a substantial shift in academic interests among college and university students in the United States, (d) an apparent improvement in the quality of entering freshman students, (e) the existence of a tighter curriculum and of higher standards in elementary and secondary schools, (f) the presence of more amiable relations between faculty and students, and (g) a projected increase in budgets for student personnel work.

A Look Ahead

It is hoped that in the next three years those undertaking studies of college personnel programs will draw heavily on theory, on methodology, and on findings represented in research recently completed and presently being conducted by sociologists, social anthropologists, and social psychologists. Such research would appear to offer much at both conceptual and practical levels.

Additional References: Becker (1961); McConnell (1962); Sanford (1962); Sprague (1960).

Bibliography

ARBUCKLE, DUGALD S., and KAUFFMAN, JOSEPH F. "Student Personnel Services in Liberal Arts Colleges." *Personnel and Guidance Journal* 38: 296-99; December 1959.

BECKERS, WILMER HENRY. *An Evaluation of the Student Personnel Program at the Auburn University School of Education.* Doctor's thesis. Auburn, Ala.: Auburn University, 1961. 374 pp. Abstract: *Dissertation Abstracts* 22: 2298-99; No. 7, 1962.

BRANTLY, EDWARD JAMES. *A Study of the Student Personnel Program at Clark College.* Doctor's thesis. Boulder: University of Colorado, 1960. 511 pp. Abstract: *Dissertation Abstracts* 21: 3372-73; No. 11, 1961.

BRUNSON, MAY A. *Guidance: An Integrating Process in Higher Education.* New York: Bureau of Publications, Teachers College, Columbia University, 1959. 180 pp.

GLANZ, EDWARD C.; HAYES, ROBERT W.; and PENNEY, JAMES F. "The Freshman Psychology Course as the Basis for a Student Personnel Program." (Comments by John W. Gustad.) *Personnel and Guidance Journal* 38: 290-95; December 1959.

GUSTAD, JOHN W. "Comments." *Personnel and Guidance Journal* 38: 294-95; December 1959.

KLITZKE, LOUIS L. "Needed Research in Junior College Student Personnel Services." *Junior College Journal* 30: 452-59; April 1960.

MCCONNELL, THOMAS R. "The Relation of Institutional Goals and Organization to the Administration of Student Personnel Work." *Approaches to the Study of Administration in Student Personnel Work, No. 9.* (Edited by Martin L. Snoke.) Minneapolis: University of Minnesota Press, 1960. Chapter 3, pp. 19-35.

MCDANIEL, J. W. *Essential Student Personnel Practices for Junior Colleges.* Washington, D.C.: American Association of Junior Colleges, 1962. 54 pp.

MEDSKER, LELAND L. *The Junior College: Progress and Prospect.* New York: McGraw-Hill Book Co., 1960. 367 pp.

MOTTO, JOSEPH J. "Development of an Extension Advisement Program." *Adult Education* 11: 235-38; Summer 1961.

MUELLER, KATE HEVNER. *Student Personnel Work in Higher Education.* Boston: Houghton Mifflin Co., 1961. 570 pp.

RICE, JOSEPH PETER, JR. *A Comparative Study of the Guidance Program in the Private and Public Junior College.* Doctor's thesis. Storrs: University of Connecticut, 1959. 180 pp. Abstract: *Dissertation Abstracts* 20: 1230-31; No. 4, 1959.

RIGGS, LAWRENCE A. *College Student Personnel Work in a Christian Context.* Nashville, Tenn.: Board of Education, Methodist Church, 1962. 89 pp.

ROBINSON, DONALD W., and BROWN, DIRCK W. "A Report on Student and Student Personnel Research Activities." *Personnel and Guidance Journal* 40: 358-60; December 1961.

SCOTT, WILLIAM L. "Student Personnel Services: Principles and Practices for Small Colleges." *Personnel and Guidance Journal* 38: 737-39; May 1960.

SIMONAITUS, JOHN. *An Investigation of the Status of the Student Personnel Program in Evening Colleges.* Doctor's thesis. Syracuse, N.Y.: Syracuse University, 1960. 142 pp. Abstract: *Dissertation Abstracts* 21: 2573; No. 3, 1961.

STARR, JAMES M. "Guidance Practices in Selected Junior Colleges in the Northwest." *Junior College Journal* 31: 442-46; April 1960.

WILLIAMSON, E. G. "An Outsider's View of Junior College Guidance Programs." *Junior College Journal* 30: 489-501; May 1960.

WILLIAMSON, E. G. *Student Personnel Services in Colleges and Universities: Some Foundations, Techniques and Processes of Program Administration.* New York: McGraw-Hill Book Co., 1961. 474 pp.

WISE, W. MAX. "Student Personnel Work—Future Trends." *Personnel and Guidance Journal* 39: 704-709; May 1961.

Additional References

BECKER, HOWARD S., and OTHERS. *Boys in White.* Chicago: University of Chicago Press, 1961. 456 pp.

MCCONNELL, THOMAS R. *A General Pattern for American Public Higher Education.* New York: McGraw-Hill Book Co., 1962. 198 pp.

SANFORD, NEVITT, editor. *The American College: A Psychological and Social Interpretation of the Higher Learning.* New York: John Wiley & Sons, 1962. 1084 pp.

SPRAGUE, HALL T., editor. *Research on College Students.* Boulder, Colo.: Western Interstate Commission for Higher Education, 1960. 188 pp.

CHAPTER IV

Selection, Preparation, and Professionalization of Specialists

ROBERT O. STRIPLING and JAMES L. LISTER *

IN THE April 1960 REVIEW, Hill and Green expressed a hope that shortages of personnel workers would result in more and better research publications on selection, preparation, and professionalization. While a review of the literature does indicate an increase in the number of publications relating to these topics, there does not seem to have been a significant increase in the reporting of basic research.

Selection of Personnel Specialists

Santavicca (1959) reported that 85 percent of the counselor education programs surveyed considered the undergraduate record of applicants. Other considerations were scholastic ability, teaching and work experience, personal adjustment, and social interests. The lack of research reported on selection may support Hill's (1961a) statement that while other factors are important, initial admission is largely contingent upon admission to graduate school. Emphasizing that counselors are drawn largely from the teaching ranks, Hill (1961b) characterized selection as a process extending from a teacher's evidenced interest in guidance work to his evaluation during a practicum experience.

Writers agreed upon the importance as well as the difficulty of assessing counselor personality characteristics. Patterson (1962) studied the characteristics of vocational rehabilitation counselor students as measured by the *Miller Analogies Test*, Kerr and Speroff's *Empathy Test*, the *Edwards Personal Preference Schedule*, the *Minnesota Multiphasic Personality Inventory*, and the *Strong Vocational Interest Blank*. The personality characteristics of National Defense Education Act (NDEA) guidance institute enrollees were studied by McDougall and Reitan (1961), who used a peer-rating technique, and by Kazienko and Neidt (1962), who differentiated between *good* and *poor* enrollees through self-descriptions on the Bennett Polydiagnostic Index. Giblette (1960) found tests of empathy, emotional stability, and practical judgment useful in differentiating among three levels of rated competency of practicing school counselors. Gazda and Ohlsen (1961) reported that prospective counselors appeared less well adjusted after seven weeks of group counseling; 14 months later, they seemed somewhat improved. Tyler (1960) described selection techniques used in the first 50 NDEA institutes. In addition to the minimum qualifications stipulated in the Act, institute staffs were encouraged to develop their own assessment techniques. Standardized tests of academic aptitude, personality, attitudes, and values were employed.

---

* Assisted by Thomas J. DiCesare.

Preparation of Counselors

Counselor preparation has received no greater impetus than that provided by the NDEA. Tyler (1960) described the guiding principles, structure, and contributions of the initial 50 institutes. Seventy-one percent of the directors surveyed by Klopf and Cohen (1962) reported that the impact of the institute programs had had an effect on their regular counselor education programs.

Riccio (1961) compared counselor-educators and state guidance supervisors who were members of the Association for Counselor Education and Supervision (ACES). He considered level of preparation, relationship between geographical area in which the degree was granted and area of current employment, degree-granting institutions, and trends in the type of doctorates earned. Polmantier and Schmidt (1960) reported a comprehensive survey of courses offered by state universities in areas designed to prepare guidance workers. Of the 54 universities surveyed, none offered courses in all 13 areas mentioned in the inquiry.

Most of the work investigating the status of preparation programs has been conducted through questionnaire surveys. Counselors who had earned master's degrees at Michigan State University were surveyed by Norris (1960). Reactions to their preparation as well as suggestions for strengthening counselor education in that institution were obtained. John Wellington (1961) surveyed 56 institutions offering courses in counselor education in the North Central Region and found that 96.4 percent offered a master's degree in counseling, with approximately 30 hours of credit. Harmon and Arnold (1960) asked members of the American School Counselor Association (ASCA) to evaluate their formal preparation. They found that 40 percent lacked supervised practice or experience in tape recording interviews and that respondents in general expressed dissatisfaction with group guidance and psychology courses. In studying programs of preparation for elementary school guidance, Hill and Nitzschke (1961) noted a lack of clarity on the part of counselor-educators concerning the functions of elementary school guidance.

Few of the counselor education programs reviewed included any systematic evaluation of behavior changes in counselor candidates. A notable exception was the study by Munger and Johnson (1960) of attitudinal changes associated with participation in a summer NDEA guidance institute. Enrollees evidenced an increase in the use of *understanding* responses and a decrease in the use of *diagnostic and probing* responses in their interactions with clients.

This period witnessed increasing debate over content areas in counselor preparation (Hoppock, 1961). One of the outstanding controversies centered around the desirability of requiring teacher certification and/or teaching experience for counselors. Hoyt (1961) and Hudson (1961) maintained that teaching experience was essential, but Cohen (1961) and Lloyd-Jones (1962) recommended experimentation to gather the needed evidence. Cohen outlined a pilot research program, and Lloyd-Jones proposed the organization of experimental counselor education programs for promising candidates from nonteaching curriculums. Sherman (1962) discussed the merits of generalist *versus* specialist approaches to counselor preparation, and Hutson (1962) recommended a teaching apprenticeship for preservice counselors. Wrenn (1962) reported that none of the 32 counselor-educators participating in the Commission on Guidance in American Schools' study mentioned teaching experience or any other particular experience when asked what the emphases would be upon counselor competencies and preparation 20 years later.

Walz and Roeber (1962) asked counselor-educators to respond to a typescript of a counseling interview. Responses indicated that supervision might be largely cognitive, information-giving, and concerned with counselor verbalization, premature evaluations, and variation in supervisory behavior. Schmidt (1960) described internship experiences at Indiana University, which included supervised practice in the schools for employed teachers as well as preservice counselors. Ivey (1961) described a summer guidance practicum at Bucknell University, which emphasized testing and test interpretation. The University of Minnesota, Counselor Education Staff (1960), described a supervised experience that included field practice in addition to a counseling practicum. The San Diego State College internship reported by Malcolm (1961) stressed the aspects of service to the schools as well as professional preparation. Secondary students who served as clients reported high satisfaction with the experience.

This period saw growing concern for the professionalization of counselors, accompanied by increasing dissatisfaction with preparation programs. These concerns were highlighted by the publication of the Wrenn (1962) report.

The Counselor in a Changing World, while written by Wrenn (1962), was an outgrowth of deliberations by the Commission on Guidance in American Schools. Presenting recommendations for counselor preparation against a background of social, cultural, scientific, and educational changes, Wrenn's work has intentionally resulted in critical discussions. Winfrey (1962) reported ASCA members' reactions to 23 recommendations from the Wrenn report and stressed the need for clarification of counselor function and preparation.

Increasing concern over standards for counselor education as well as the need for clarification of counselor role, functions, and conditions of work led to the launching of the studies by the Association for Counselor Education and Supervision and American School Counselor Association (1962). Involving regional chairmen and local study committees, the goals of these national studies are to provide standards of preparation and descriptions of counselor role, functions, and conditions of work that will ultimately receive endorsement by the parent organization, the American Personnel and Guidance Association (APGA). The interim report of the Association for Counselor Education and Supervision and American School Counselor Association (1962) contains position and reaction papers presented at the 1962 APGA convention. In addition to these major efforts by committees and organizations, such individuals as Hitchcock (1962), Dugan (1961), McCully (1962), Feder (1961), and Pierson and Grant (1959) emphasized the need for decisive steps toward professionalization.

Schmidt (1962) studied counselor role perceptions held by school counselors and their principals. Using a Q-sort technique, he compared inter-counselor and principal-counselor agreement on actual and ideal counselor functions. While many areas of significant agreement were noted, the observed relationships, which failed to account for much of the existing variation, suggested that areas of substantial disagreement remained.

Professionalization of Counselors in Educational Settings

Walker (1961) reported the efforts made by Akron, Ohio, schools to alleviate their shortage of professionally qualified counselors through in-service education. Jean Wellington (1961) described efforts to involve graduate students from Tufts University in the program of the Winchester, Massachusetts, guidance department. Camp (1959) described a systematic on-the-job education program in the Baltimore County schools that was strongly supported by school personnel. A questionnaire survey by Buonagura (1960) of 32 New Jersey junior high schools revealed that one-third of the approved schools were actually conducting in-service education programs. Of the approaches described, only the faculty meeting was used in all schools. Evaluations of these programs were highly subjective.

A reliable index of increased professionalization of any group of specialists is its movement to adopt certification standards. The latest U.S. Office of Education report (Brewster, 1960) listed 37 states with mandatory certification requirements and 4 with optional requirements, compared with 34 mandatory and 7 optional, respectively, in the previous report (Brewster, 1957). While certification requirements have increased, it should be noted that specific requirements vary greatly even among states; furthermore, one may question, in the absence of evidence, how many employed counselors actually satisfy these minimum requirements. Also, it should be noted that the essential question concerning quality of preparation has not been examined. Lloyd (1962) compared attitudes of counselors and counselor-educators toward the mandatory certification requirements of 34 states. Both groups expressed dissatisfaction with certain aspects of existing requirements as well as disagreement regarding desirable certification. This led Lloyd to note the need for research on the adequacy of expressed opinions as well as their most effective implementation.

While the number of states with mandatory certification requirements is increasing, such standards may not really provide a satisfactory basis for quality control (Dugan, 1962). Dugan further noted that reciprocity among states was impossible until more common bases for certification were nationally recognized. He recommended the initiation of a national cooperative study of standards. McCully (1961) questioned whether the school counselor's functions were sufficiently unique to justify certification; he cited the absence of certification requirements in some states as evidence that the functions performed by the school counselor lack sufficient recognition to warrant nationwide certification.

Concern for ethical behavior of personnel workers was evidenced at the 1961 APGA convention by the adoption of a code of ethics as set forth by Arsenian (1961). The adoption of this code may well be another major step in the direction of professionalization of counselors.

Although the legal position of counselors remains unclear, Garber (1960) stated that the Iverson case is a step in the direction of needed clarification. Miller and Simpson (1961) presented a review of the legal aspects of the counseling relationship under the headings of expert witness, professional immunity, and counselor liability.

While the exact ratio of counselors to students in the public schools remains a controversial issue, few doubt the existence of a serious shortage of counselors. Crummel (1961) estimated that there is a shortage of more than 18,000. Dugan (1961) estimated that at least 5,000 qualified new counselors will be needed each year through 1970 and suggested selecting prospective counselors from a variety of academic backgrounds.

Many factors seem to contribute to the wide range of salaries. Wrenn (1962) gave a median salary of approximately $6,000 for the counselors in Project Talent. In this sample, 9 percent earned less than $4,000 and 10 percent, over $8,000. It should be noted that 47 percent of these counselors were employed for school terms of 10 months.

Preparation and Professionalization of Other Specialists

Little research appropriate for this chapter has been reported concerning the preparation and professionalization of specialists other than counselors. The scarcity of publications considering such specialists as the school psychologist, social worker, and health worker from the broad *pupil personnel* point of view may mean that such a point of view has not gained widespread acceptance yet.

In considering personnel specialists in broad perspective, one finds a minimum of evidence on the incidence of such specialists in the public schools. An exception was a statistical report of the American Nurses' Association (1961), which showed approximately 10,000 nurses employed in schools.

Berg and Berg (1962) have been explicit in differentiating between the roles of the school nurse and the teacher in the area of school health. They emphasized that while their respective roles are distinctive, they are both members of the school health team. Klein (1959), who was more explicit about the guidance functions of the nurse, pointed out that health counseling was given in 82 percent of the public schools surveyed. She also reported that organized guidance programs existed in about one-half of the public schools employing nurses, and that, in these schools, nurses were usually considered members of the guidance team.

A statement on functions and qualifications of school nurses adopted by the American Nurses' Association, Public Health Nurses, School Nurses Branch (1961), stressed understanding of school and community as well as mastery of specific nursing skills. Specific functions relating to counseling and guidance were directed toward the elimination or minimizing of health problems that interfere with learning and development.

In discussing the social worker's function in the school, Nebo (1960) emphasized the importance of an interdisciplinary and interprofessional approach involving a broad array of school and community personnel.

Summary

Research on selection, preparation, and professionalization of personnel specialists remains limited largely to questionnaire surveys and position papers. While the current debates are inevitable and desirable concomitants of professional growth, one would hope to see them soon paralleled by basic research.

Bibliography

American Nurses' Association. *Facts About Nursing: A Statistical Summary*. New York: the Association (10 Columbus Circle), 1961. 255 pp.

American Nurses' Association, Public Health Nurses, School Nurses Branch. "Functions and Qualifications for School Nurses." *American Journal of Nursing* 61: 93-97; May 1961.

Arsenian, Seth. "The A.P.G.A. Code of Ethics." *Personnel and Guidance Journal* 40: 204-209; October 1961.

Association for Counselor Education and Supervision and American School Counselor Association. *Counselor Education: A Progress Report on Standards*. Washington, D.C.: American Personnel and Guidance Association, 1962. 90 pp.

Berg, Edgar E., and Berg, Marie P. "The Teacher's and the Nurse's Roles in School Health." *Nursing Outlook* 10: 190-92; March 1962.

Brewster, Royce E. *Guidance Workers Certification Requirements*. U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1957, No. 22. Washington, D.C.: Government Printing Office, 1957. 58 pp.

Brewster, Royce E. *Guidance Workers Certification Requirements.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1960, No. 14. Washington, D.C.: Government Printing Office, 1960. 98 pp.

Buonaguro, Paul Vincent. *In-Service Guidance Training for New Jersey Junior High School Teachers.* Doctor's thesis. New Brunswick, N.J.: Rutgers University, 1960. 120 pp. Abstract: *Dissertation Abstracts* 21: 1840-41; No. 7, 1961.

Camp, N. Harry, Jr. "A County Studies Its In-Service Counselor Training Program." *Personnel and Guidance Journal* 38: 309-13; December 1959.

Cohen, Nancy K. "Must Teaching Be a Prerequisite for Guidance?" *Counselor Education and Supervision* 1: 69-71; Winter 1961.

Crummel, Robert A. "Analysis of Standards for Guidance Programs Under NDEA." *Social Life* 44: 22-24; November 1961.

Dugan, Willie E. "Critical Concerns of Counselor Education." *Counselor Education and Supervision* 0: 5-11; Spring 1961.

Dugan, Willie E. "An Inward Look: Assumptions and Aspirations." *Counselor Education and Supervision* 1: 174-80; Summer 1962.

Feder, Daniel D. "Perspectives and Challenges." *Personnel and Guidance Journal* 40: 6-10; September 1961.

Garber, Lee O. "Wisconsin Court Defines Legal Responsibilities of School Guidance Counselors." *Nation's Schools* 65: 78; January 1960.

Gazda, George, and Ohlsen, Merle. "The Effects of Short Term Group Counseling on Prospective Counselors." *Personnel and Guidance Journal* 39: 634-38; April 1961.

Giblette, John Franklin. *Differences Among Above Average, Average, and Below Average Secondary School Counselors.* Doctor's thesis. Philadelphia: University of Pennsylvania, 1960. 146 pp. Abstract: *Dissertation Abstracts* 21: 812-13; No. 4, 1960.

Harmon, Donald, and Arnold, Dwight L. "High School Counselors Evaluate Their Formal Preparation." *Personnel and Guidance Journal* 39: 303-306; December 1960.

Hill, George E. "The Counselor Educator's Relationship with Counselors, School Administrators, and Teachers." *Counselor Education and Supervision* 0: 16-23; Spring 1961. (a)

Hill, George E. "The Selection of School Counselors." *Personnel and Guidance Journal* 39: 355-60; January 1961. (b)

Hill, George E., and Nitzschke, Dale F. "Preparation Programs in Elementary School Guidance." *Personnel and Guidance Journal* 40: 155-59; October 1961.

Hitchcock, Arthur A. "From the Executive Director." *Personnel and Guidance Journal* 40: 498-503; January 1962.

Hoppock, Robert. "Is Psychology Enough?" *Counselor Education and Supervision* 0: 28; Spring 1961.

Hoyt, Kenneth B. "What the School Has a Right To Expect of Its Counselors." *Personnel and Guidance Journal* 40: 129-32; October 1961.

Hudson, George R. "Counselors Need Teaching Experience." *Counselor Education and Supervision* 0: 24-27; Spring 1961.

Hutson, P. W. "Another 'Position' Paper." *Counselor Education and Supervision* 2: 40-44; Fall 1962.

Ivey, Allen E. "A Summer Guidance Practicum." *Counselor Education and Supervision* 1: 48-51; Fall 1961.

Kazienko, Louis W., and Neidt, Charles O. "Self Descriptions of Good and Poor Counselor Trainees." *Counselor Education and Supervision* 1: 106-23; Spring 1962.

Klein, Ruth A. "The School Nurse as a Guidance Functionary." *Personnel and Guidance Journal* 38: 318-21; December 1959.

Klopf, Gordon, and Cohen, Nancy K. "The Impact of the NDEA Counseling and Guidance Institutes on the Professional Education of School Counselors." *Counselor Education and Supervision* 1: 151-61; Spring 1962.

Lloyd, David O. "Counselor and Counselor Trainer Attitudes Toward Counselor Certification in the United States." *Personnel and Guidance Journal* 40: 791-98; May 1962.

Lloyd-Jones, Esther. "Implications of the Wrenn Report for Counselor Education." *Counselor Education and Supervision* 2: 17-25; Fall 1962.

McCully, C. Harold. "A Rationale for Counselor Certification." *Counselor Education and Supervision* 1: 3-9; Fall 1961.

McCully, C. Harold. "School Counselor: Strategy for Professionalization." *Personnel and Guidance Journal* 40: 681-89; April 1962.

McDougall, William P., and Reitan, Henry M. "The Use of a Peer Rating Technique in Appraising Selected Attributes of Counselor Trainees." *Counselor Education and Supervision* 1: 72-76; Winter 1961.

Malcolm, David D. "San Diego State College Counseling Internship: A 'Team' Approach." *Personnel and Guidance Journal* 39: 382-87; January 1961.

Miller, Frank W., and Simpson, Richard J. "Some Legal Aspects of the Counselor Client Relationship." *Counselor Education and Supervision* 1: 19-29; Fall 1961.

Munger, Paul F., and Johnson, Carletona. "Changes in Attitude Associated with an NDEA Counseling and Guidance Institute." *Personnel and Guidance Journal* 38: 751-53; May 1960.

Nebo, John C., editor. *Administration of School Social Work.* New York: National Association of Social Workers, 1960. 64 pp.

Norris, Willa. "More than a Decade of Training Guidance and Personnel Workers." *Personnel and Guidance Journal* 39: 287-91; December 1960.

Patterson, C. H. "Test Characteristics of Rehabilitation Counselor Trainees." *Journal of Rehabilitation* 28: 15-16; September-October 1962.

Pierson, George A., and Grant, Claude W. "The Road Ahead for the School Counselor." *Personnel and Guidance Journal* 38: 207-10; November 1959.

Polmantier, Paul C., and Schmidt, Lyle D. "Areas of Preparation for School Guidance Workers." *Personnel and Guidance Journal* 39: 45-46; September 1960.

Riccio, Anthony C. "The Counselor Educator and the Guidance Supervisor: Graduate Training and Occupational Mobility." *Counselor Education and Supervision* 1: 10-17; Fall 1961.

Santavicca, G. Gene. "Supervised Experience and Selection of Counselor Trainees." *Personnel and Guidance Journal* 38: 195-97; November 1959.

Schmidt, Louis G. "Guidance Internship Training at Indiana University." *Personnel and Guidance Journal* 38: 577-79; March 1960.

Schmidt, Lyle D. "Concepts of the Role of Secondary School Counselors." *Personnel and Guidance Journal* 40: 601-605; March 1962.

Sherman, Robert. "The School Counselor: Generalist or Specialist?" *Counselor Education and Supervision* 1: 203-11; Summer 1962.

Tyler, Leona E. *The National Defense Counseling and Guidance Training Institutes Program: A Report of the First 50 Institutes.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1960, No. 31. Washington, D.C.: Government Printing Office, 1960. 80 pp.

University of Minnesota, Counselor Education Staff. "Supervised Field Practice in Student Personnel Work." *Personnel and Guidance Journal* 39: 308-309; December 1960.

Walker, Robert N. "A Counselor Development Program Meets a Guidance Manpower Need." *Personnel and Guidance Journal* 39: 472-75; February 1961.

Walz, Garry R., and Roeber, Edward C. "Supervisor's Reactions to a Counseling Interview." *Counselor Education and Supervision* 2: 2-7; Fall 1962.

Wellington, Jean. "Observation Program: An Informal Experiment in Counselor Training in the Public Schools." *Personnel and Guidance Journal* 40: 176-77; October 1961.

Wellington, John A. "Concepts of Counselor Training." *Counselor Education and Supervision* 0: 29-33; Spring 1961.

Winfrey, James K. "Implications of the Wrenn Report to the School Counselor." *Counselor Education and Supervision* 2: 35-39; Fall 1962.

Wrenn, C. Gilbert. *The Counselor in a Changing World.* Washington, D.C.: American Personnel and Guidance Association, 1962. 195 pp.

CHAPTER V

Counseling

ROBERT CALLIS

THE RESEARCH on counseling reported in the three-year period covered by this REVIEW has been extensive and meaningful. With few exceptions, the methodology employed was sophisticated and appropriate for the task, and the researchers seemed to be genuinely interested in discovering new knowledge rather than in proving theoretical convictions.

Theory Construction

It was heartening to see several articles in which attempts were made to synthesize and to integrate theories pertaining to counseling. In Callis' (1960) article, the key concept was that clients who seek assistance from counselors have some current or anticipated inadequacy in their behavior repertoire. Inadequacies could be caused by (a) lack of experience, (b) distorted perceptions, and (c) erroneous generalizations. The task of counseling is to discover inadequacies and ways of correcting them. These discoveries can be made by either the client or the counselor or by both. Two general methods of counseling were postulated from (a) client self-discovery and (b) counselor discovery and the interpretation of the discovery to the client.

In Layton's (1961) model, the principal ingredients were (a) external reality, (b) perceptions of self and others, and (c) communication between reality and the perceived self. Perceptions and possible distortion in perceptions are central concepts. One of the goals of counseling is to bring the client's self-perception into closer congruence with reality. It would not be difficult to blend the notions of Callis and Layton into a single theoretical model. Also, many of the notions of self-theorists could be encompassed in this model. It would appear then that the two constructs—perception and external reality—are becoming embedded firmly in the theoretical formulations about counseling.

Additional Reference: Weitz (1961).

Avowed Goals

As has been either stated or implied in the previously described theoretical formulations, the avowed goal of counseling is to assist the client in changing his behavior repertoire so that he can behave in a more effective and personally satisfying manner. Tyler (1960), in differentiating counseling from psychotherapy, stated that the goal of counseling is to bring about better utilization of the personal resources that the client already possesses. She indicated that the object of counseling is not a major reorganization of personality but a minimum change in psychological direction. To accomplish this, it may be necessary to reduce the client's anxiety and thereby to increase the client's problem-solving behavior.

Berdie (1959) indicated that the goal of counseling is behavioral change. He suggested not a sweeping personality change but the solving of immediate, specific problems, which, in and of itself, may be sufficient to remobilize the client. He agreed with Tyler when he indicated that the reduction of anxiety and of defense mechanisms may be a necessary prelude to increasing the problem-solving skills of the client. Berdie cited some behavioral changes that may be the specific goals of specific clients, e.g., higher grades in school or college, longer persistence in an education program (fewer dropouts), more accurate understanding of self, and greater self-acceptance. Apostal (1960) and Gonyea (1962) proposed appropriateness of occupation engaged in (or planned for) as an outcome criterion for certain cases.

Additional References: Curran (1960); Goodstein and Grigg (1959); Samler (1960).

Obtained Results

Educational

Goodstein and Crites (1961) studied the effect of brief counseling on grades earned by low-ability college freshmen. The experimental group was given brief counseling; the control group was not. Difference in grade point average earned by the two groups favored the control group for the summer session (the experimental period) and for the following semester. Richardson (1960) conducted a similar study and found that the counseled and noncounseled students did not differ significantly in earned grade point average. In both studies, the students in the experimental group were invited by the counseling service to be counseled. In both studies, the students were of low academic ability compared with the total freshman class. Richardson concluded that the student must be motivated to come for help before counseling can produce favorable results.

Searles (1962) conducted a study similar to the two cited above, but used superior college freshmen as subjects. The experimental group members were afforded three interviews each during the fall semester; the control group members were afforded none. The interviews, conducted by the director of student activities of the college, consisted primarily of informing the student about his academic ability and his academic progress. At the end of the semester, the two groups were compared on several academic criteria. There were no significant differences between the two groups. Searles reported that many of the students in the experimental group expressed no interest in or need for the counseling they received.

Baymur and Patterson (1960) identified 32 underachieving high school juniors and assigned them to four matched groups of eight students each. Group A received individual counseling, 10-12 interviews per student. Group B received group counseling for 9 sessions. Group C received a one-session group experience in which the students were informed that they were underachievers and were exhorted to do something about it. Group D received no treatment. Criterion data, obtained at the beginning and at the end of the experimental period of one semester, were (a) a measurement of personal adjustment, (b) a measure of study habits and attitudes, and (c) the change in grade point average between the experimental semester and the previous semester. The four groups did not differ significantly on any of the criteria. However, a comparison of the combined counseled groups with the combined noncounseled groups indicated that the counseled groups made significant gains in personal adjustment and grade point average.

Watson (1961) studied the effect of counseling on college students who had been referred for counseling because of personal-emotional problems. The 66 students in Group A accepted the referral and made use of counseling (a median of 10 interviews); the 45 students in Group B denied any need for help to the intake interviewer; the 23 students in Group C made no response to the offer of counseling. The three groups were approximately equal in academic ability and in academic performance at the time of referral. Two years later, 68 percent of Group A, 51 percent of Group B. and 43 percent of Group C had either been graduated from or were still enrolled in a degree program. However, differences in motivation must be held, along with counseling, as a contributing factor to the observed criterion difference.

Shepherd (1962) used graduation from college as the criterion for counseling. The experimental group consisted of 295 students who had requested and received counseling at some time during their college career. The control group consisted of classmates of the experimental group matched by sex, academic ability, and academic division. Five years after the subjects had enrolled as freshmen, Shepherd ascertained that about 50 percent of the counseled group had been graduated, while about 37 percent of the control group had been graduated. There was no significant difference in academic ability between the graduates and the nongraduates. As in the case of Watson's (1961) study, possible motivational differences between the two groups were not controlled.

Vocational

Apostal (1960) compared the judgments of counseling adequacy based on case-folder data with the judged appropriateness of occupation engaged in approximately seven years after counseling and found no relationship between these two criteria. Gonyea (1962) found that a significant increase in appropriateness of vocational choice had occurred between the time that counseling had begun and terminated. However, there was some question about whether this increase was still in effect a year after counseling had terminated.

Lurie, Goldfein, and Baxt (1960) studied the effect of vocational counseling on a group of underachieving ninth graders. The experimental group was provided individual counseling and extensive group activities with heavy emphasis on orientation to the occupational world. At the end of a year, they found that the experimental group showed more improvement in appropriateness of occupational plans and had a lower dropout rate than did a no-treatment matched control group.

Williams (1962) used different criteria for evaluating the outcomes of brief educational-vocational counseling, an index of personal adjustment. and an index of congruence between actual and ideal self. He used three groups—an experimental group who received *typical* vocational counseling of approximately two interviews conducted by an experienced counselor; a client-control group who had requested counseling, but were placed on a waiting list for the duration of the experimental period; and a nonclient group. He found that counseling restored a normal level of personal adjustment and degree of congruence for the experimental group. However, no change was observed in the control groups. These improvements were in effect a year after counseling had terminated.

Additional Reference: Thompson (1960).

Personal-Social-Emotional

Davis (1959) studied the effect of group guidance and individual counseling on citizenship behavior. From a group of 70 high school seniors. Davis identified the 30 with the lowest citizenship grades for the previous year and assigned these 30 at random to three groups. Members of Group A were provided 20 periods of group guidance; Group B, 2 individual counseling interviews; and Group C, no treatment. Group C showed no improvement in citizenship grades during the experimental period. Groups A and B showed significant improvement, with Group A exhibiting the most improvement. Each subject in Group A, of course, received 10 times as much counselor time as did each subject in Group B. But the intriguing aspect of this comparison was that the counselor himself spent an equal amount of time on Group A and on Group B.

Kemp (1961) studied the effect of dogmatism (closed-mindedness *versus* open-mindedness) on counseling. Two experimental groups (high and low dogmatism) were given four individual counseling interviews during a 10-week period. The two control groups (high and low dogmatism) received no treatment. Neither of the control groups made any significant change in score on a problem check list. The high-dogmatic experimental group made no change in scores on the problem check list, but the low-dogmatic experimental group did. Again, a forced counseling situation exists, but in this study one may have a clue to what kind of person will profit from counseling under such conditions.

Arbuckle and Boy (1961) studied the effect of client-centered counseling on high school students with behavior problems. Thirty-six students who had been placed on after-school disciplinary detention were divided into three matched groups. Members of Group A were released from detention on the condition that they report for counseling for a 12-week period; Group B, continued on detention and received no counseling; Group C, released from detention and received no counseling. At the end of the experimental period, Group A showed an increase in congruence between ideal self and actual self and a favorable increase in teachers' behavior ratings. Groups B and C showed no such change.

Additional References: Baker (1960); Hunt and others (1959); Zax and Klein (1960).

Diagnosis

In most of the studies cited in this chapter, crude but obvious diagnostic classification schemes have been employed, often grudgingly. Tyler (1960. 1961) reluctantly decided that diagnosis is necessary, but preferred to use a different label. Arbuckle (1961) frankly rejected the whole idea of diagnosis in counseling. Apostal and Miller (1959) described a two-dimensional diagnostic classification plan, which was subsequently used by Shepherd (1962) and by Kirk (1962). Shepherd found that the graduation rate of clients varied sharply, depending on the diagnosis. Kirk found that the problems presented to high school counselors by clients differed only in minor respects from those presented to college counselors. As pointed out by Apostal (1960), there is a clear danger of masking significant results if different kinds of counseling cases are not studied separately.

Additional References: Kirk (1959); Weiner (1959).

Methods

Wrenn (1960) studied the effect of theoretical conviction on counselor responses and found that the only difference associated with theoretical orientation was that psychoanalytically oriented counselors used reflection less frequently than other counselors. On the contrary, the principal variations in counselor behavior were associated with the topic being discussed. Brams (1961) found that the only personality characteristic of counseling trainees related to their ability to communicate with clients was a tolerance for ambiguity.

Hoffman (1959) developed a system for specifying and classifying counselor interview behavior by identifying several roles that counselors use in interviews, such as teaching and listening.

Frank and Sweetland (1962) studied the effects of different kinds of counselor response on subsequent client response. Clarification of feeling responses was followed by client responses of understanding and insight. Counselor interpretation also gave rise to an increase in understanding and insight responses. These findings raise serious questions about earlier formulations that suggested that the reflection and clarification of feeling are the only counselor responses that can produce client understanding and insight.

The extensive literature on laboratory studies dealing with how one person can influence the behavior of another has been summarized by Krasner (1958), Schachter (1959), and Thibaut and Kelley (1959). McDavid (1959) found that group-oriented people were influenced to change more often than were task-oriented people. Staples and Walters (1961) noted that firstborns were more acquiescent and suggestable. Masling (1960) cited many studies indicating that the personality or behavior of the examiner significantly influenced the nature of projective test responses given by the subject.

Russell (1961) coached his clients so that some were negative and some positive in their approaches to counselors. The negative clients generated more counselor anxiety than did the positive ones.

Martin, Lundy, and Lewin (1960) had one group of subjects talk to a tape recorder while no one else was present. Another group talked to a counselor who remained silent, and a third group talked to a counselor who responded. Compared with the other two groups, the third group evidenced greater approach behavior to emotional material and a greater decrease in galvanic skin response reactivity. Quay (1959) instructed interviewers to reinforce family memories for one group of subjects and nonfamily memories for another group. The reinforcement was found to be effective in eliciting the desired responses. These and many other studies clearly indicate that under certain conditions one person can influence the verbal behavior of another in an intended direction.

Additional References: Dinoff and others (1960); Grigg (1961); Heilbrun (1961); Ivey (1962); Pohlman and Robinson (1960); Pumroy (1961); Waskow (1962).

Summary

During the past three years, progress has been made in clarifying the domain and goal of counseling and in developing theories to explain the process. Although the evidence is not conclusive, it strongly suggests that unless clients perceive a need for assistance from a counselor, the counselor is apt to have little impact on client behavior. Trends in the data indicate that experienced counselors may produce better results than neophytes and that most experienced counselors choose their methods to fit the task rather than a stereotyped school of conviction. Laboratory studies of counseling prototypes clearly indicate that it is possible for one person to influence the behavior of another in an intended direction.

Considerable clarification is still needed in the employment of outcome criteria. In many instances, the outcome against which counseling was evaluated was not the intended goal of the counselor but an incidental by-product. Knowledge of definitive cause-and-effect relationships in the counseling process would be greatly increased if research data were sorted into homogeneous classifications by some kind of diagnostic formulations applied to client and counselor characteristics and behavior.

Bibliography

Apostal, Robert A. "Two Methods of Evaluating Vocational Counseling." *Journal of Counseling Psychology* 7: 171-75; Fall 1960.

Apostal, Robert A., and Miller, John G. *A Manual for the Use of a Set of Diagnostic Categories.* University of Missouri Testing and Counseling Service Report, Vol. 13, No. 4. Columbia: the University, 1959. 9 pp.

Arbuckle, Dugald S. *Counseling: An Introduction.* Boston: Allyn and Bacon, 1961. 349 pp.

Arbuckle, Dugald S., and Boy, Angelo V. "Client-Centered Therapy in Counseling Students with Behavior Problems." *Journal of Counseling Psychology* 8: 136-39; Summer 1961.

Baymur, Feriha B., and Patterson, Cecil H. "A Comparison of Three Methods of Assisting Underachieving High School Students." *Journal of Counseling Psychology* 7: 83-89; Summer 1960.

Berdie, Ralph F. "Counseling Principles and Presumptions." *Journal of Counseling Psychology* 6: 175-82; Fall 1959.

Brams, Jerome M. "Counselor Characteristics and Effective Communication in Counseling." *Journal of Counseling Psychology* 8: 25-30; Spring 1961.

Callis, Robert. "Toward an Integrated Theory of Counseling." *Journal of College Student Personnel* 1: 2-9; June 1960.

Davis, Donald A. "Effect of Group Guidance and Individual Counseling on Citizenship Behavior." *Personnel and Guidance Journal* 38: 142-45; October 1959.

Frank, George H., and Sweetland, Anders. "A Study of the Process of Psychotherapy: The Verbal Intersection." *Journal of Consulting Psychology* 26: 135-38; April 1962.

Gonyea, George G. "Appropriateness-of-Vocational-Choice as a Criterion of Counseling Outcome." *Journal of Counseling Psychology* 9: 213-19; Fall 1962.

Goodstein, Leonard D., and Crites, John O. "Brief Counseling with Poor College Risks." *Journal of Counseling Psychology* 8: 318-21; Winter 1961.

Hoffman, A. Edward. "An Analysis of Counselor Sub-Roles." *Journal of Counseling Psychology* 6: 61-67; Spring 1959.

Kemp, C. Gratton. "Influence of Dogmatism on Counseling." *Personnel and Guidance Journal* 39: 662-65; April 1961.

Kirk, David M. *Diagnostic Classification of Problems Presented by High School Counselees.* Doctor's thesis. Columbia: University of Missouri, 1962. 99 pp.

Krasner, Leonard. "Studies of the Conditioning of Verbal Behavior." *Psychological Bulletin* 55: 148-70; May 1958.

Layton, Wilbur L. "Constructs and Communication in Counseling: A Limited Theory." *Journal of Counseling Psychology* 8: 3-7; Spring 1961.

Lurie, Walter A.; Goldfein, Jacob; and Baxt, Roland. "An Intensive Vocational Counseling Program for Slow Learners in High School." *Personnel and Guidance Journal* 39: 21-29; September 1960.

McDavid, John W., Jr. "Personality and Situational Determinants of Conformity." *Journal of Abnormal and Social Psychology* 58: 241-46; March 1959.

Martin, Barclay; Lundy, Richard M.; and Lewin, Mark H. "Verbal and GSR Responses in Experimental Interviews as a Function of Three Degrees of 'Therapist' Communication." *Journal of Abnormal and Social Psychology* 60: 234-40; March 1960.

Masling, Joseph. "The Influence of Situational and Interpersonal Variables in Projective Testing." *Psychological Bulletin* 57: 65-85; January 1960.

Quay, Herbert. "The Effect of Verbal Reinforcement on the Recall of Early Memories." *Journal of Abnormal and Social Psychology* 59: 254-57; September 1959.

Richardson, LaVange H. "Counseling the Ambitious Mediocre Student." *Journal of Counseling Psychology* 7: 265-68; Winter 1960.

Russell, Peter D. *Counselor Anxiety in Relation to Clinical Experience and Hostile or Friendly Clients.* Doctor's thesis. University Park: Pennsylvania State University, 1961. 206 pp. Abstract: *Dissertation Abstracts* 23: 327-28; No. 1, 1962.

Schachter, Stanley. *The Psychology of Affiliation.* Stanford, Calif.: Stanford University Press, 1959. 141 pp.

Searles, Aysel, Jr. "The Effectiveness of Limited Counseling in Improving the Academic Achievement of Superior College Freshmen." *Personnel and Guidance Journal* 40: 630-33; March 1962.

Shepherd, Robert E. *An Evaluation of Counseling Using the Criterion of College Graduation.* Doctor's thesis. Columbia: University of Missouri, 1962. 81 pp.

Staples, Frederick R., and Walters, Richard H. "Anxiety, Birth Order and Susceptibility to Social Influence." *Journal of Abnormal and Social Psychology* 62: 716-19; May 1961.

Thibaut, John W., and Kelley, Harold H. *The Social Psychology of Groups.* New York: John Wiley & Sons, 1959. 313 pp.

Tyler, Leona E. "Minimum Change Therapy." *Personnel and Guidance Journal* 38: 475-79; February 1960.

Tyler, Leona E. *The Work of the Counselor.* Second edition. New York: Appleton-Century-Crofts, 1961. 327 pp.

Watson, Gladys H. "An Evaluation of Counseling with College Students." *Journal of Counseling Psychology* 8: 99-104; Summer 1961.

Williams, John E. "Changes in Self and Other Perceptions Following Brief Educational-Vocational Counseling." *Journal of Counseling Psychology* 9: 18-30; Spring 1962.

Wrenn, Robert L. "Counselor Orientation: Theoretical or Situational?" *Journal of Counseling Psychology* 7: 40-45; Spring 1960.

Additional References

Baker, Elliott. "The Differential Effects of Two Psychotherapeutic Approaches on Client Perceptions." *Journal of Counseling Psychology* 7: 46-50; Spring 1960.

Curran, Charles A. "Some Ethical and Scientific Values in the Counseling Psychotherapeutic Process." *Personnel and Guidance Journal* 39: 15-20; September 1960.

Dinoff, Michael, and others. "An Experimental Analogue of Three Psychotherapeutic Approaches." *Journal of Clinical Psychology* 16: 70-73; January 1960.

Goodstein, Leonard D., and Grigg, Austin E. "Client Satisfaction, Counselors, and the Counseling Process." *Personnel and Guidance Journal* 38: 19-24; September 1959.

Grigg, Austin E. "Client Response to Counselors at Different Levels of Experience." *Journal of Counseling Psychology* 8: 217-22; Fall 1961.

Heilbrun, Alfred B., Jr. "Male and Female Personality Correlates of Early Termination in Counseling." *Journal of Counseling Psychology* 8: 31-36; Spring 1961.

Hunt, J. McV., and others. "An Integrated Approach to Research on Therapeutic Counseling with Samples of Results." *Journal of Counseling Psychology* 6: 46-54; Spring 1959.

IVEY, ALLEN E. "Role Conflict in Counseling: Its Effect on College Student Attitudes." *Journal of Counseling Psychology* 9: 139-43; Summer 1962.

KIRK, BARBARA A. "Counseling Graduate Students." *Journal of Counseling Psychology* 6: 284-87; Winter 1959.

POHLMAN, EDWARD, and ROBINSON, FRANCIS P. "Client Reaction to Some Aspects of the Counseling Situation." *Personnel and Guidance Journal* 38: 546-51; March 1960.

PUMROY, DONALD K. "Some Counseling Behavior Correlates of the Social Desirability Scale." *Journal of Counseling Psychology* 8: 49-53; Spring 1961.

SAMLER, JOSEPH. "Changes in Values: A Goal in Counseling." *Journal of Counseling Psychology* 7: 32-39; Spring 1960.

THOMPSON, ALBERT S. "Personality Dynamics and Vocational Counseling." *Personnel and Guidance Journal* 38: 350-57; January 1960.

WASKOW, IRENE E. "Reinforcement in a Therapy-Like Situation Through Selective Responding to Feelings or Content." *Journal of Consulting Psychology* 26: 11-19; February 1962.

WEINER, IRVING B. "The Role of Diagnosis in a University Counseling Center." *Journal of Counseling Psychology* 6: 110-15; Summer 1959.

WEITZ, HENRY. "Guidance as Behavior Change." *Personnel and Guidance Journal* 39: 550-60; March 1961.

ZAX, MELVIN, and KLEIN, ARMIN. "Measurement of Personality and Behavior Changes Following Psychotherapy." *Psychological Bulletin* 57: 435-48; September 1960.

CHAPTER VI

The Appraisal Function

LEO GOLDMAN

THIS CHAPTER is restricted to publications that were concerned with or particularly relevant to the appraisal function itself. Excluded are most studies regarding the reliabilities, validities, and norms of tests. Although several of these studies contain important substantive and methodological contributions in the area of appraisal, many of them have been reviewed in the February 1962 REVIEW, *Educational and Psychological Testing*. Although a somewhat more inclusive approach was taken with nontest methods of appraisal, only a very few items were located.

General treatments of the appraisal function, including references to relevant research, were found in Cottle and Downie (1960) and in Goldman (1961). The three issues of the *Annual Review of Psychology* (1960, 1961, 1962) that almost correspond to the period covered by this issue of the REVIEW had no separate chapters on counseling, although some pertinent material was included under other chapter headings.

The various reviews continued to reflect pessimism regarding the validities of tests and of other sources of appraisal data. In particular, there was criticism of the lack of theoretical foundations for much of the work with tests.

Test Taking and Its Preparation

Instructions

Hills (1961) compared the results on alternate forms of the *California Test of Personality* when subjects were told to answer rapidly at one administration and deliberately at another. Although differences were found, it was not clear to what extent they might have been associated with a lack of comparability in the factorial composition of the alternate forms of the inventory.

Practice

In their study of practice effects in relation to spatial ability tests, Krumboltz and Christal (1960) discovered that alternate forms administered up to seven hours apart yielded as much apparent practice effect as did a retest with the same form. It was concluded that in order for practice effects to be minimized, use of different types of spatial tests with unlike specific variance would be necessary.

Perception

The subject's perceptions of a test and of its significance would be expected to influence his set, the effort he expended in responding to the items, and, therefore, his scores. Wrightsman (1960) found no significant differences in average score between those subjects who had been told that the test was a valid measure of intelligence and those subjects who had been told that it was not valid. Moreover, there was no difference between the two groups in the extent to which they took advantage of an opportunity to cheat through using not-well-erased answers in the test booklets. It would be interesting to know how individual subjects really perceived the test, as an unimportant laboratory exercise or as a test that was measuring something important to their self-concept.

Screening

Several investigators reported on the effectiveness of personality instruments in the screening of high school students (Billard, 1962) and college students (Caldwell, 1959; Goodstein and others, 1961; Kleinmuntz, 1960, 1961; Parker, 1961a, b). There was general agreement that the personality profile at the time of counseling reflected more disturbance than the one taken during the orientation period. There was disagreement concerning the identifiability at screening time of those who later would seek counseling. To some extent, this circumstance might stem from the different kinds of students that different counseling centers were attracting.

Scoring and Score Conversions

Burack (1961) noted serious errors in machine scoring of the *Strong Vocational Interest Blank* (*SVIB*). When seven men's *SVIB* answer sheets were scored a second time by the same machine scoring agency, 35 percent of the occupational scores changed by at least one score level. Rescoring of two women's *SVIB* answer sheets yielded changes on 52 percent of the scales. Even after the agency reported that the machines had been adjusted, 22 percent of the scores on a single answer sheet changed on rescoring.

The importance of local norms was reinforced in several studies. Although Goodstein and others (1961) found that the *California Psychological Inventory* differentiated among counseling center clients and other students on a campus, the scores of the latter group departed substantially from the published norms for college students. An interchange between

Kleinmuntz (1960, 1961) and Parker (1961a, b) led to the conclusion that different cutting scores for the *Minnesota Multiphasic Personality Inventory* were needed at two different universities on a scale of maladjustment.

Nontest Methods

As usual, there were very few reports of research with nontest methods of appraisal. The gap between possible and actual practice was seen in Hutson's (1960) summary of several dissertations on character-trait ratings. Several of these studies showed that carefully designed, properly used rating scales yielded relatively high correlations between such traits as reliability and industry and high school grades and between these same traits and college grades. On the other hand, a study of the rating practices of 100 high schools judged to be among the best in the United States led to the conclusion that most such plans violated the sound principles of rating that were established in the 1920's.

Weiss and Dawis (1960) studied the accuracy of information obtained in interviews with 48 physically handicapped clients of the Division of Vocational Rehabilitation and with 43 Employment Service applicants. When agency records and employer reports were used as criteria, the accuracy of the information varied from 100 percent for sex identification to 45 percent for the question on whether the person had received any assistance from the agency. It should be noted that the interviews were conducted not by counselors but by opinion-polling interviewers who apparently had visited the homes. The results might have been different if the information had been obtained during counseling interviews.

Interpreting Data

The largest number of studies relating to the appraisal function touched on the interpretation of data—that process by which inferences are made about a person or group.

Expectancy Studies

Although no major methodological research was located, several studies reflected a small increase in the use of established methods of producing specific expectancy data. One of the most ambitious projects was a statewide program for relating high school grades and college aptitude test scores to college grades (Hills and others, 1961). The product was a manual for school counselors, which provided tables of norms and of regression equations for each of the cooperating colleges and universities in Georgia.

The College Entrance Examination Board (1963) made available to counselors a valuable compendium of normative data about the freshman classes at 250 member colleges. The presentations were not all uniform, and not all colleges furnished complete information. The data included distributions of scores on the *Scholastic Aptitude Test* of the College Entrance Examination Board and of high school rank of applicants as well as distributions on the same variables for accepted students and for scholarship winners. Despite its limitations, this kind of information should be a boon to high school counselors who have the ability and understanding necessary to use it properly.

One of the major avenues of improvement in data interpretation has been the attention given to determination of local norms. Kaczkowski (1959), for example, described methods of preparing local expectancy tables. Juola (1961) used college freshmen data to illustrate several kinds of expectancy tables, several of which showed the interaction of two or three predictor variables and one criterion variable. One high school followed up its graduates into college and prepared distributions of high school rank for each college (Leussler, 1961). Four small schools collaborated in preparing expectancy tables that demonstrated the relationship between scores on an aptitude test for engineering and physical science and grades in an engineering college within a university (Oliver, 1960).

Counselor Versus Statistical Prediction

A single scale on the *SVIB* predicted the job satisfaction of engineers with far greater accuracy than did the judgment of any one of six college counselors, who seemed to rely on the use of one or more other scales of the *SVIB* (Lewis and MacKinney, 1961). The *SVIB* Group IV scale, which is considered suitable for technicians, correlated .50 with the reported job satisfaction of engineers. The counselors seemed to weight the scale used for engineers more heavily than the one for technicians. Brayfield's (1961) comment on the article suggested an explanation for the outcomes of the investigation. Subjects who were low in occupational self-concept (the Group IV scale typically correlates negatively with occupational level) were more satisfied with the occupational status that they attained as engineers than were those individuals whose expectations had been higher. If this hypothesis is true, it would appear that the failure of these counselors, as that of so many other counselors in studies of clinical *versus* statistical methods, stemmed from inadequate understanding and conceptualization of the worker-situation interaction.

Strong Vocational Interest Blank

Stephenson (1961a) developed a method for classifying groups of scales on the *SVIB* that appeared to be more reliable than the Darley-Hagenah method. A chart that listed all possible combinations of letter grades on the scales in each group served to reduce the number of cases in which arbitrary classifications could be made.

Stephenson (1961b, 1962) also investigated the chance scores or gray areas on the *SVIB*. The first study demonstrated that gray-area scores were similar in retest reliability to other scores. Therefore, Stephenson concluded that gray-area scores were not likely to be the result of chance, but rather a consequence of low interest of the subjects relative to people in the occupation. The second study was less convincing. Subjects taking the *SVIB* on a retest were asked to fake in the direction of obtaining higher scores or lower scores as well as to fake in obtaining chance scores for a stated occupation. Subjects could fake successfully in attaining higher or lower scores but not in achieving chance scores. One questions the possibility of the task assigned the fake-chance subjects. A person can fake in the direction of a high or low score by visualizing the likes and dislikes of people in an occupation, but how does he visualize the chance responses of people in an occupation?

Filbeck and Callis (1961) developed a verification scale for the *SVIB* that included 103 rare item-responses. The scale differentiated genuine from theoretical chance profiles on a cross-validation study. Reliability and normative data were also presented, as was the scale itself.

Brown (1961) pointed out that a common error in interpretation of the *SVIB* results from misunderstanding the expectancy ratio. Strong's conclusion that chances are 3.6 to 1 that a man with an A rating on an occupational scale will enter a specific occupation is applicable only to samples that are similar to Strong's sample in terms of the percentages of men who later enter each occupation in question.

Some long-range validity data led Berdie (1960) to suggest that careful use of the *SVIB* with high school seniors is justified.

Differential Predictability

Two studies shed further light on individual differences in the predictability of criteria from the same predictor variables. Berdie (1961) found some evidence, though inconclusive, that a person's variability of response to the items on an ability test was negatively related to the predictability of his grades in college. Ghiselli (1960) continued his contributions in this area by describing procedures for determining which of two valid measures is the better predictor for a given individual.

These two studies continued a line of research which may prove exceedingly valuable in both theoretical and applied outcomes. Not only may one expect improvement in the predictive efficiency of test and other data, but also one may gain increased understanding of the dynamic interaction of various personal and environmental factors.

Miscellaneous

Katz (1962) explained that a profile on the *Kuder Preference Record* (*KPR*) is the result of both ipsative (intra-individual) and normative (interindividual) comparisons. He concluded that since one cannot legitimately evaluate a *KPR* profile in either ipsative or normative terms, the interpretability of this inventory is open to question.

Knox and Hoffman (1962) found that the graphic dimensions of a test profile influenced interpretation by college students. In the analysis of data obtained, broader variance in judgments, higher reliability, and greater coefficient of multiple correlation were associated with the presence of the relatively wide spread of scores on the percentile graph than with the appearance of a relatively narrow range of scores on the standard-score graph.

Discrepancy scores between amount of detail on the *Draw-a-Man* test and the *Ohio State University Psychological Test* (*OSUPT*) did produce a higher degree of prediction of scholastic achievement for counseled probation students than did the *OSUPT* alone (Isard and Lasky, 1961). The authors suggested that the discrepancy score reflects the extent of integration among a person's different abilities.

Communicating Appraisal Information

Although several related studies appeared, little research was reported which dealt directly with the process of communicating appraisal information to individuals or groups. This lack of published research constitutes an apparent decrease from earlier years.

From a county school system in Florida, Durost (1959) reported that locally based stanine scores were found meaningful to teachers and parents. Burr (1959) described how the IBM card on which the scores were profiled represented an effective use of test results. Eells (1961) and Harris and Dole (1960) presented interesting graphic methods for portraying expectancy data.

Those teachers who had made least use of guidance records rated most highly a procedure for transmitting to them a summary of data taken from the cumulative records (Hopke, 1961).

Evaluation of Test Reports

Again, short-term test-reporting programs were found to effect little or no change in student estimates of their abilities and of their anticipated grades (Robertson, 1960a, b). On the other hand, Herman and Zeigler (1961) reported that some degree of change was realized with parents of college students. It is questionable that these studies tapped deep-seated or lasting changes. Furthermore, it would be surprising if such changes did occur as the result of one individual or group session.

Conclusions

The last few years saw a small burst of published research on the appraisal function, including such topics as analysis of clinical interpretation processes, comparison of clinical and statistical methods of data interpretation, the influence of different approaches to selecting tests in the counseling process, and the effectiveness of different methods of reporting appraisal data and of making interpretations. Perhaps one is seeing merely another example of what is too frequently an occurrence in educational research: abandonment of a topic after a few studies have been reported. Yet the earlier studies on the appraisal function surely did not exhaust the subject; they merely suggested intriguing hypotheses that often challenged counselors to justify their questionable appraisal activities.

Bibliography

Annual Review of Psychology. Palo Alto, Calif.: Annual Reviews. Vol. 11, 1960, 544 pp. Vol. 12, 1961, 533 pp. Vol. 13, 1962, 602 pp.

BERDIE, RALPH F. "Strong Vocational Interest Blank Scores of High School Seniors and Their Later Occupational Entry." *Journal of Applied Psychology* 44: 161-65; June 1960.

BERDIE, RALPH F. "Intra-Individual Variability and Predictability." *Educational and Psychological Measurement* 21: 663-76; Autumn 1961.

BILLARD, RICHARD G. "Comparison of Teacher Ratings of Personality with Results of an Incomplete Sentences Blank." *Personnel and Guidance Journal* 41: 58-59; September 1962.

BRAYFIELD, ARTHUR H. "Comments." *Journal of Counseling Psychology* 8: 229-30; Fall 1961.

BROWN, FRED G. "A Note on Expectancy Ratios, Base Rates, and the SVIB." *Journal of Counseling Psychology* 8: 368-69; Winter 1961.

BURACK, BENJAMIN. "Have You Checked Machine-Scoring Error Lately?" *Vocational Guidance Quarterly* 9: 191-93; Spring 1961.

BURR, WILLIAM L. "Increasing the Effective Use of Test Results by Machine Processing." *Sixteenth Yearbook: National Council on Measurements Used in Education.* (Edited by Edith M. Huddleston.) New York: the Council, 1959. pp. 149-54.

CALDWELL, EDWARD. "Stability of Scores on a Personality Inventory Administered During College Orientation Week." *Personnel and Guidance Journal* 38: 305-308; December 1959.

COLLEGE ENTRANCE EXAMINATION BOARD. *Manual of Freshman Class Profiles.* New York: the Board, 1963. 600 pp.

COTTLE, WILLIAM C., and DOWNIE, NORVILLE M. *Procedures and Preparation for Counseling.* Englewood Cliffs, N.J.: Prentice-Hall, 1960. 330 pp.

DUROST, WALTER N. "The Use of Local Stanines in Reporting Test Results in a Large Cosmopolitan School System." *Sixteenth Yearbook: National Council on Measurements Used in Education.* (Edited by Edith M. Huddleston.) New York: the Council, 1959. pp. 140-48.

EELLS, KENNETH. "A Vivid Method of Presenting Chances for Academic Success." *Journal of Counseling Psychology* 8: 344-50; Winter 1961.

FILBECK, ROBERT W., and CALLIS, ROBERT. "A Verification Scale for the Strong Vocational Interest Blank, Men's Form." *Journal of Applied Psychology* 45: 318-24; October 1961.

GHISELLI, EDWIN E. "Differentiation of Tests in Terms of the Accuracy with Which They Predict for a Given Individual." *Educational and Psychological Measurement* 20: 675-84; Winter 1960.

GOLDMAN, LEO. *Using Tests in Counseling.* New York: Appleton-Century-Crofts, 1961. 434 pp.

GOODSTEIN, LEONARD D., and OTHERS. "The Use of the California Psychological Inventory in a University Counseling Service." *Journal of Counseling Psychology* 8: 147-53; Summer 1961.

HARRIS, YEUELL Y., and DOLE, ARTHUR A. "A Pilot Study in Local Research with the Differential Aptitude Test Battery." *Personnel and Guidance Journal* 39: 128-32; October 1960.

HERMAN, LOUIS M., and ZEIGLER, MARTIN L. "The Effectiveness of Interpreting Freshman Counseling-Test Scores to Parents in a Group Situation." *Personnel and Guidance Journal* 40: 143-49; October 1961.

HILLS, JOHN R. "The Influence of Instructions on Personality Inventory Scores." *Journal of Counseling Psychology* 8: 43-48; Spring 1961.

HILLS, JOHN R., and OTHERS. "Admissions and Guidance Research in the University System of Georgia." *Personnel and Guidance Journal* 39: 452-57; February 1961.

HOPKE, WILLIAM. "Getting Guidance Information into the Hands of Teachers." *School Counselor* 9: 62-65; December 1961.

HUTSON, P. W. "Recent Studies in Character-Trait Rating." *Personnel and Guidance Journal* 38: 364-68; January 1960.

ISARD, ELEANORE S., and LASKY, DAVID I. "A Discrepancy Score Method in Predicting Scholastic Achievement of College Freshmen Counseled During Probation." *Personnel and Guidance Journal* 39: 725-28; May 1961.

JUOLA, ARVO E. "Multi-Variable Grade Expectancy Tables: An Aid to Test Interpretation." *Eighteenth Yearbook: National Council on Measurement in Education.* (Edited by Edith M. Huddleston.) Ames, Iowa: the Council, 1961. pp. 91-99.

KACZKOWSKI, HENRY R. "Using Expectancy Tables To Validate Test Procedures in High School." *Educational and Psychological Measurement* 19: 675-77; Winter 1959.

KATZ, MARTIN. "Interpreting Kuder Preference Record Scores: Ipsative or Normative." *Vocational Guidance Quarterly* 10: 96-100; Winter 1962.

KLEINMUNTZ, BENJAMIN. "Identification of Maladjusted College Students." *Journal of Counseling Psychology* 7: 209-11; Fall 1960.

KLEINMUNTZ, BENJAMIN. "Screening: Identification or Prediction?" *Journal of Counseling Psychology* 8: 279-80; Fall 1961.

KNOX, ROBERT E., and HOFFMAN, PAUL J. "Effects of Variation of Profile Format on Intelligence and Sociability Judgments." *Journal of Applied Psychology* 46: 14-20; February 1962.

KRUMBOLTZ, JOHN D., and CHRISTAL, RAYMOND E. "Short-Term Practice Effects in Tests of Spatial Aptitude." *Personnel and Guidance Journal* 38: 385-91; January 1960.

LEUSSLER, DONN. "Hints for Counselors: An Exciting Follow-Up Study." *School Counselor* 9: 17-18; October 1961.

LEWIS, EDWIN C., and MACKINNEY, ARTHUR C. "Counselor vs. Statistical Predictions of Job Satisfaction in Engineering." (Comment by Arthur H. Brayfield.) *Journal of Counseling Psychology* 8: 224-30; Fall 1961.

OLIVER, GARY L. "Cooperative Guidance Research Among Smaller Schools." *Vocational Guidance Quarterly* 8: 202-204; Summer 1960.

PARKER, CLYDE A. "Letter to the Editor." *Journal of Counseling Psychology* 8: 88-89; Spring 1961. (a)

PARKER, CLYDE A. "The Predictive Use of the MMPI in a College Counseling Center." *Journal of Counseling Psychology* 8: 154-58; Summer 1961. (b)

ROBERTSON, MALCOLM H. "Counselor and Student Estimates of Grades as Predictors of Academic Achievement." *Journal of Educational Research* 54: 73-75; October 1960. (a)

ROBERTSON, MALCOLM H. "Test Scores and Self-Estimates of Two Curricular Groups." *Personnel and Guidance Journal* 38: 746-50; May 1960. (b)

STEPHENSON, RICHARD R. "Chance Versus Nonchance Scores on the SVIB." *Journal of Applied Psychology* 45: 415-19; December 1961. (a)

STEPHENSON, RICHARD R. "A New Pattern Analysis Technique for the SVIB." *Journal of Counseling Psychology* 8: 355-62; Winter 1961. (b)

STEPHENSON, RICHARD R. "Faking 'Chance' on the SVIB." *Journal of Applied Psychology* 46: 252-56; August 1962.

WEISS, DAVID J., and DAWIS, RENE V. "An Objective Validation of Factual Interview Data." *Journal of Applied Psychology* 44: 381-85; December 1960.

WRIGHTSMAN, LAWRENCE S., JR. "The Effects of the Purported Validity of a Test upon Level of Motivation and Performance." *Journal of Educational Research* 54: 153-56; December 1960.

CHAPTER VII

Vocational Development Process

GARRY R. WALZ

VOCATIONAL development as a major area of investigation is young and vigorous. If one traces its beginnings to Super's Career Pattern Study, it is scarcely more than a decade old. In that short span, investigations of vocational behavior have moved from conceptually shallow trait-and-factor approaches to fertile occupational models that encompass the behavioral sciences.

Although more than 200 separate sources were easily identified for this chapter, publications on vocational behavior have some of the qualities of an iceberg. Out of sight are the many studies and activities that provide the buoyancy for those that are published. The author has chosen to be objective rather than impartial, but much of interest and worth has been excluded. Throughout, articles of major theoretical or research importance have been favored.

Career Development

Some of the characteristics of the newer conceptual approaches given by Borow (1961) in his succinct analysis of research on vocational development were (a) production of more hypothetical constructs, (b) use of both psychodynamics and ego psychology, and (c) preference for longitudinal research designs. He calls convincingly for less research on the prediction of vocational choice and more on vocational striving and the meaning of work. In another review of research, Super, Tiedeman, and Borow (1961) identified three issues to be resolved: (a) development of adequate statistical tools for career prediction, (b) construction of both theoretical and operational definitions of vocational behavior, and (c) development of comparable measures of vocational maturity at different age levels.

A major empirical test of some of the hypotheses from Super's theories was contained in an investigation of the vocational maturity of ninth grade boys by Super and Overstreet (1960). Using six indexes of vocational maturity, they defined vocational maturity in the ninth grade as primarily preparation for choice and planning ahead. Super (1961) also found that the preferences of ninth grade boys were inadequate as criteria for determining the need for or the effectiveness of counseling. Among Super's significant contributions are his continuing insightful theorizing and the testing of his hypotheses through a longitudinal career development study.

Roe's theory linking childhood experiences to occupational interests was the subject of three investigations. Hagen's (1960) results did not generally support Roe's posited relationship between family experiences and later career choice. In a study of ministerial and chemistry students. Switzer and others (1962) found little in the students' childhoods that was predictive of occupational choice. In an investigation of four groups of professional women, Utton (1962) found that only one Roe-based hypothesis was supported: social workers and occupational therapists manifested a greater altruistic love of people than did dietitians and laboratory technicians. Generally, the studies did not support Roe's hypotheses. It is hard to generalize, however, since restricted samples and retrospective approaches were used. Perhaps Roe's major contribution to date has been to offer a series of hypotheses about family influences in career development, which have been a spur to research.

Two studies were related to Holland's personality theory of vocational development. Holland (1960) obtained empirical support for his belief that preferences for occupational titles are significantly related to a number of personality variables. Tentative support for Holland's contention that a person's level of occupational choice and aspiration is indicative of his evaluation of himself was obtained by Schutz and Blocher (1961). Holland's aim to develop an integrating theory of vocational choice that is explicit enough to encourage research would seem to be a desirable and attainable goal.

Operating on her premise that the core of individuality consists of a person's choices and the manner in which he organizes them, Tyler (1961) investigated subjects' choices, using methods from interest-measurement research and concept-formation research. Her intent was to discover the individuality inherent in people who make common job choices. Progress to date has been encouraging; further validity studies lie ahead.

Using Bordin's theory of occupational choice, Schutz and Blocher (1960) found a significant relationship between vocational preferences and occupational stereotypes selected as self-descriptive by twelfth grade boys. Blocher and Schutz (1961) found that the twelfth grade students perceived both their self-concepts and their ideal concepts to be closer to stereotypes of workers in occupations with high claimed interests than to their equivalent stereotypes of workers in occupations with little claimed interests.

Notable research contributions were made by Tiedeman and his co-workers. Their work encompassed formulation of a psychosocial model. development of new statistical techniques, and practical application of their findings. An investigation into the clarification of boys' self-concepts was undertaken by O'Hara and Tiedeman (1959). Their results would seem to demonstrate that boys' self-concepts about work values and interests were clarified as they moved from grade 9 to grade 12. Aptitudes, however, were found to be relatively poorly perceived, even by able students, in grades 9-12. It was hypothesized that there is an interest stage ending by grade 10 and a work-values stage extending through grade 12. Cass and Tiedeman (1960) found that interests were effective predictors of the selection of high school curriculum and that aptitudes not only were poor predictors of choice of curriculum but also were not even well understood until the senior year of high school.

One of the most encouraging aspects of career development research was the use of theoretically rich models as a basis of empirical investigations. Hopefully, within three or four years, one will see the use of broader populations and more longitudinal designs. If one can judge a theory by the research it fosters, it would seem that the work of Super, Roe, and Tiedeman could be regarded as major contributions. Bordin, Holland, and Tyler also served as sources for considerable research activity.

Additional References: Borow (1960); Davis, Hagan, and Strouf (1962); Flanagan and others (1962); Holland (1961); Kibrick and Tiedeman (1961); Nachmann (1960); Segal (1961); Stephenson (1961); Stern (1961); Super (1960); Tiedeman (1961); Uzzell (1961).

Occupational Psychology

In perhaps the first long-term longitudinal study of job satisfaction, Hoppock (1960) conducted a 27-year follow-up of employed adults. He found older persons to be more satisfied, with the largest increases in job satisfaction occurring among those who changed jobs. Useful as his study was, additional and more adequately controlled studies need to be made before one can generalize.

Three studies were concerned with the relationship of personality needs to occupations. Dipboye and Anderson (1961) found that high school seniors have occupational stereotypes that can be expressed in terms of manifest needs. Nurturance, endurance, and order were perceived as needs satisfied by the largest number of occupations. Gonyea (1961) used Andrews' *A* technique to identify the dimensions by which college students view occupations. The usual interest and job classification grouping did not appear. A relation between the personality needs of individuals and the liking or disliking of job duties individuals are called upon to perform as members of a broad occupational group was discovered by Walsh (1959).

In an investigation of the personality characteristics of occupational groups, Miller (1962) found that the groups differed on personality variables that appeared to be relevant to the characteristics of the occupations. The variables could be divided into those dealing with goals and those related to means.

Kinnane and Suziedelis (1962), using Super's Work Values Inventory, found a high degree of relationship between work values and inventoried interests. It was hypothesized that an individual's patterning of interests reflects his values regarding work. In a study of the ordering of occupational values by high school students, Dipboye and Anderson (1959) found a tendency for all age and sex groups to rank interesting work first and security second. Independence was ranked low by all student groups. It appeared that occupational values were formed by the ninth grade and that they did not change during the remainder of high school. Both studies highlighted the need for further research on development of occupational values and the meaning of variances in values between and within occupational groups.

Nosow and Form (1962) edited an excellent book of readings on the sociology of work. Personality as it relates to occupational choice and job performance was a subject of an increasing number of investigations. Definitive research on the perceptions of occupations by people both within and outside an occupation are needed. There is room, too, for research on what the need-meeting qualities of occupations are and on how a worker tailors his job to meet his personal needs.

Additional References: Crites (1962); Cureton and Katzell (1962); Froehlich and Wolins (1960); Harding and Bottenberg (1961); Katzell, Barrett, and Parker (1961); Robinson and Connors (1960, 1961); Vroom (1962).

Occupational Classification

Few developments were noted in expanding upon the early promise of the classification systems introduced in the April 1960 issue of the REVIEW. Norton (1962) studied the judgments of 70 workers about the needed capacities and education for jobs with the ratings given in the *Worker Trait Requirements*. He concluded that while the ratings needed refinements, they would still be useful to a counselor.

In a critical appraisal of the *Functional Occupational Classification Project* (*FOCP*), Walther (1960) demonstrated that important worker characteristics were not identified by use of the system proposed in the *FOCP*. He argued for a classification structure based on differences between people rather than differences in work situations. An occupational system was proposed by Brender (1960) whereby the intrinsic job satisfaction in each occupation would be identified and described.

The counseling use of present classification systems calls for considerable counselor adaptation if emphasis is given to the meaning of a job to a worker. More useful systems may have to wait for further research in career development.

Women's Career Development

White (1959) investigated the relationships of self-concept and of parental identification to girls' vocational interests and occupational choices. Compared with girls seeking careers, those aspiring to traditionally feminine roles were more satisfied with themselves and came from more stable homes. Moreover, the girls with traditional leanings did furnish evidence of clearer communication with their parents than did the career-oriented girls. The post high school plans of senior girls were found by Milliken (1961) to be realistically related to scholastic aptitude.

A novel approach to the study of women's interests was undertaken by Warren (1959), who sought to relate the measured interests of college women to occupational role. She found that the *Strong Vocational Interest Blank* did differentiate among the occupational roles held by women. More than 60 percent had high interest scores in occupations other than housewife and elementary teacher.

Girls today are faced with conflicting pressures in developing an occupationally relevant self-concept. There is also evidence that the majority of women will devote an increasingly large share of their adult lives to work outside the home. Yet little is known about the vocational development of women. The previously described studies contributed needed, albeit limited, information. Restricted populations and cross-sectional designs are hardly sufficient for the task. It would seem that the time is now right for longitudinal developmental studies of girls similar to Super's study of boys.

Sources of Occupational Information

A major challenge to the field is the development of occupational materials that emphasize career development. Most of the materials smack of an earlier job choice era. A notable exception is the *Widening Occupational Roles Kit* by Lifton (1962), which seeks to promote the vocational maturity of junior high school students. Fictionalized job briefs and filmstrips are used to assist the student in occupational planning and decision making.

Two studies on readability provide little optimism for improvement in the reading level of occupational materials. Using both the Flesch and the Dale-Chall readability formulas, Watson, Rundquist, and Cottle (1959) concluded that little had occurred in the last decade to make occupational materials more readable. The reading levels of most materials were inappropriate for junior high school students; the materials on professional and semiprofessional occupations were especially difficult. Samples of occupational materials studied by Diener and Kaczkowski (1960-61) showed a wide range of readability on a given occupation between publishers and between occupations within a given publication source. Until occupational materials are written at the lowest reading level of potential users, counselors will have to check for readability.

In a particularly provocative article, Samler (1961) suggested that printed occupational information was concerned primarily with the "economic man"—with wages, working conditions, and other factors established by job analysis. Consideration of the "psychological man"—the worker's role and the way his personality needs will be met—was practically nonexistent. Support was given for a research program to assess individual needs and to discover the need-related characteristics of a given work situation.

Knowledge seems to be at hand to improve occupational materials. Some changes are being made. Used as a stimulus for self-exploration, occupational materials offer potentialities not yet realized. Research on the career-development use of occupational materials could provide the framework for the development of more appropriate materials.

Summary

A new ferment was evident in the area described by the term *vocational development*. There was also a notable lag. Neither materials nor practices showed the impact of the newer theoretical formulations. Relatively few investigators have done much to change the theoretical thrust. It will take counselors as a group to implement the newer conceptions. Only by the counselors' actions will those materials and procedures be developed that capitalize on the mounting knowledge of vocational behavior.

Bibliography

BLOCHER, DONALD H., and SCHUTZ, RICHARD A. "Relationships Among Self-Descriptions, Occupational Stereotypes, and Vocational Preferences." *Journal of Counseling Psychology* 8: 314-17; Winter 1961.

BOROW, HENRY. "Vocational Development Research: Some Problems of Logical and Experimental Form." *Personnel and Guidance Journal* 40: 21-24; September 1961.

BRENDER, MYRON. "Toward a Psychodynamic System of Occupational Counseling." *Journal of Counseling Psychology* 7: 96-100; Summer 1960.

CASS, JOHN C., and TIEDEMAN, DAVID V. "Vocational Development and the Election of a High School Curriculum." *Personnel and Guidance Journal* 38: 538-45; March 1960.

DIENER, THOMAS F., and KACZKOWSKI, HENRY R. "Readability of Occupational Information." *Vocational Guidance Quarterly* 9: 87-89; Winter 1960-61.

DIPBOYE, W. J., and ANDERSON, W. F. "The Ordering of Occupational Values by High School Freshmen and Seniors." *Personnel and Guidance Journal* 38: 121-24; October 1959.

DIPBOYE, W. J., and ANDERSON, W. F. "Occupational Stereotypes and Manifest Needs of High School Students." *Journal of Counseling Psychology* 8: 296-304; Winter 1961.

GONYEA, GEORGE G. "Dimensions of Job Perceptions." *Journal of Counseling Psychology* 8: 305-12; Winter 1961.

HAGEN, DOUGLAS. "Careers and Family Atmospheres: An Empirical Test of Roe's Theory." *Journal of Counseling Psychology* 7: 251-56; Winter 1960.

HOLLAND, JOHN L. "The Relation of the Vocational Preference Inventory to the Sixteen Personality Factor Questionnaire." *Journal of Applied Psychology* 44: 291-96; August 1960.

HOPPOCK, ROBERT. "A Twenty-Seven Year Follow-Up on Job Satisfaction of Employed Adults." *Personnel and Guidance Journal* 38: 489-92; February 1960.

KINNANE, JOHN F., and SUZIEDELIS, ANTANAS. "Work Value Orientation and Inventoried Interests." *Journal of Counseling Psychology* 9: 144-48; Summer 1962.

LIFTON, WALTER M. *Widening Occupational Roles Kit.* Chicago: Science Research Associates, 1962.

MILLER, SUTHERLAND, JR. "Relationship of Personality to Occupation, Setting, and Function." *Journal of Counseling Psychology* 9: 115-21; Summer 1962.

MILLIKEN, ROBERT. "Post-High School Plans of Senior Girls in Relation to Scholastic Aptitude." *Vocational Guidance Quarterly* 10: 49-52; Autumn 1961.

NORTON, JOSEPH L. "A Study of the Estimates of Worker Trait Requirements." *Personnel and Guidance Journal* 41: 59-61; September 1962.

NOSOW, SIGMUND, and FORM, WILLIAM, editors. *Man, Work, and Society.* New York: Basic Books, 1962. 612 pp.

O'HARA, ROBERT P., and TIEDEMAN, DAVID V. "Vocational Self-Concept in Adolescence." *Journal of Counseling Psychology* 6: 292-301; Winter 1959.

SAMLER, JOSEPH. "Psycho-Social Aspects of Work: A Critique of Occupational Information." *Personnel and Guidance Journal* 39: 458-65; February 1961.

SCHUTZ, RICHARD A. and BLOCHER, DONALD H. "Self-Concepts and Stereotypes of Vocational Preferences." *Vocational Guidance Quarterly* 8: 241-44; Summer 1960.

SCHUTZ, RICHARD A., and BLOCHER, DONALD H. "Self-Satisfaction and Level of Occupational Choice." *Personnel and Guidance Journal* 39: 595-98; March 1961.

SUPER, DONALD E. "Consistency and Wisdom of Vocational Preferences as Indices of Vocational Maturity in the Ninth Grade." *Journal of Educational Psychology* 52: 35-43; February 1961.

SUPER, DONALD E., and OVERSTREET, PHOEBE L. *The Vocational Maturity of Ninth Grade Boys.* New York: Bureau of Publications, Teachers College, Columbia University, 1960. 212 pp.

SUPER, DONALD E.; TIEDEMAN, DAVID V.; and BOROW, HENRY. "Some Unresolved Issues in Vocational Development Research." *Personnel and Guidance Journal* 40: 11-14; September 1961.

SWITZER, DAVID K., and OTHERS. "Early Experiences and Occupational Choice: A Test of Roe's Hypothesis." *Journal of Counseling Psychology* 9: 45-48; Spring 1962.

TYLER, LEONA E. "Research Explorations in the Realm of Choice." *Journal of Counseling Psychology* 8: 195-201; Fall 1961.

UTTON, ALDEN C. "Recalled Parent-Child Relations as Determinants of Vocational Choice." *Journal of Counseling Psychology* 9: 49-53; Spring 1962.

WALSH, RICHARD P. "The Effect of Needs on Responses to Job Duties." *Journal of Counseling Psychology* 6: 194-98; Fall 1959.

WALTHER, REGIS H. "The Functional Occupational Classification Project: A Critical Appraisal." *Personnel and Guidance Journal* 38: 698-706; May 1960.

WARREN, PHYLLIS A. "Vocational Interests and the Occupational Adjustment of College Women." *Journal of Counseling Psychology* 6: 140-47; Summer 1959.

WATSON, DONALD E.; RUNDQUIST, RICHARD M.; and COTTLE, WILLIAM C. "What's Wrong with Occupational Materials?" *Journal of Counseling Psychology* 6: 288-91; Winter 1959.

WHITE, BECKY J. "The Relationship of Self Concept and Parental Identification to Women's Vocational Interests." *Journal of Counseling Psychology* 6: 202-206; Fall 1959.

Additional References

BOROW, HENRY. "Research Programs in Career Development." *Journal of Counseling Psychology* 7: 62-70; Spring 1960.

CRITES, JOHN O. "An Interpersonal Relations Scale for Occupational Groups." *Journal of Applied Psychology* 46: 87-90; April 1962.

CURETON, EDWARD E., and KATZELL, RAYMOND A. "A Further Analysis of the Relations Among Job Performance and Situational Variables." *Journal of Applied Psychology* 46: 230; June 1962.

DAVIS, DONALD A.; HAGAN, NELLIE; and STROUF, JUDIE. "Occupational Choice of Twelve-Year-Olds." *Personnel and Guidance Journal* 40: 628-29; March 1962.

FLANAGAN, JOHN C., and OTHERS. *Design for a Study of American Youth.* Boston: Houghton Mifflin Co., 1962. 240 pp.

FROEHLICH, HERBERT P., and WOLINS, L. "Job Satisfaction as Need Satisfaction." *Personnel Psychology* 13: 407-20; Winter 1960.

HARDING, FRANCIS D., and BOTTENBERG, ROBERT A. "Effect of Personal Characteristics on Relationships Between Attitudes and Job Performance." *Journal of Applied Psychology* 45: 428-30; December 1961.

HOLLAND, JOHN L. "Some Explorations with Occupational Titles." (Comment by Henry Borow.) *Journal of Counseling Psychology* 8: 82-87; Spring 1961.

KATZELL, RAYMOND A.; BARRETT, RICHARD S.; and PARKER, TREADWAY C. "Job Satisfaction, Job Performance, and Situational Characteristics." *Journal of Applied Psychology* 45: 65-72; April 1961.

KIBRICK, ANNE K., and TIEDEMAN, DAVID V. "Conception of Self and Perception in Schools of Nursing." *Journal of Counseling Psychology* 8: 62-69; Spring 1961.

NACHMANN, BARBARA. "Childhood Experience and Vocational Choice in Law, Dentistry, and Social Work." *Journal of Counseling Psychology* 7: 243-50; Winter 1960.

ROBINSON, H. ALAN, and CONNORS, RALPH P. "Job Satisfaction Researches of 1959." *Personnel and Guidance Journal* 39: 47-52; September 1960.

ROBINSON, H. ALAN, and CONNORS, RALPH P. "Job Satisfaction Researches of 1960." *Personnel and Guidance Journal* 40: 373-77; December 1961.

SEGAL, STANLEY J. "A Psychoanalytic Analysis of Personality Factors in Vocational Choice." *Journal of Counseling Psychology* 8: 202-10; Fall 1961.

STEPHENSON, RICHARD R. "Occupational Choice as a Crystallized Self-Concept." *Journal of Counseling Psychology* 8: 211-16; Fall 1961.

STERN, H. H. "A Follow-Up Study of Adolescents' Views of Their Personal and Vocational Future." *British Journal of Educational Psychology* 31: 170-82; June 1961.

SUPER, DONALD E. "The Critical Ninth Grade: Vocational Choice or Vocational Exploration." *Personnel and Guidance Journal* 39: 106-109; October 1960.

TIEDEMAN, DAVID V. "Decision and Vocational Development: A Paradigm and Its Implications." *Personnel and Guidance Journal* 40: 15-21; September 1961.

UZZELL, ODELL. "Influencers of Occupational Choice." *Personnel and Guidance Journal* 39: 666-69; April 1961.

VROOM, VICTOR H. "Ego-Involvement, Job Satisfaction and Job Performance." *Personnel Psychology* 15: 159-77; Summer 1962.

CHAPTER VIII

Group Procedures

E. WAYNE WRIGHT *

MUCH AKIN to the Biblical account of the tower of Babel is the "confusion of tongues" surrounding group procedures in guidance. In fact, one may question whether much change has occurred in this regard since the publication of the April 1960 REVIEW, in which Hoyt and Moore pointed up the same semantic problems and suggested that, for the defensiveness of opposing camps, the proverbial lamb of individual counseling might well lie down with the multiple counseling lion. However, a few mediating attempts in the literature of the past three years are encouraging. Especially pertinent are the contributions of Goldman (1962) and of Mahler and Caldwell (1961), who proposed meaningful differences in process and in group composition as bases for distinguishing between group guidance, group counseling, and group therapy. It was Goldman's opinion that unsuccesssful experiences with group methods of counseling and guidance have been due largely to lack of understanding of the roles of content and process in the group activity being undertaken.

Wrenn (1962) gave some support to Goldman's premise by suggesting that the counselor of the future be aware of the values of various group learning experiences and that he define, on a continuum somewhere between group guidance in the classroom and group psychotherapy, the particular province of the counselor in utilizing planned group experiences as a supplement to individual counseling. Rinn (1961) pointed up the need for clarifying the distinctive roles and purposes of different group leaders and indicated that instructional *versus* developmental-centered orientations might well represent mutually exclusive determinants of separate group processes. He concluded that the type of process employed would dictate the nature of the group leader's attitudes and his subsequent role in the group.

The topic discussions that follow reflect several other trends noted in the literature. The headings encompass those aspects of group work in guidance that received attention in the literature during the past three years.

Group Counseling

Group counseling received, by far, the greatest attention of all group guidance activities. In the studies reviewed, the use of the term *group*

---

* The author wishes to acknowledge the assistance of Glendon Casto, Karen Brown, and Jean Heaps.

counseling exceeded significantly the references to *multiple counseling*, although the terms were generally considered interchangeable as distinguishable from *group guidance*. At the same time, studies by Davis (1959) and by Kemp (1962), which reported experimental comparisons between group guidance and individual counseling, were not clear in specifying the type of group activity under investigation. The reader could not be sure, in either study, whether the group guidance process being evaluated against individual counseling was really multiple counseling, which certainly points up the need for continued efforts toward a resolution of the semantic problem mentioned previously.

The attention given to group counseling was accompanied by greater acceptance of this procedure as a valid concept of counseling per se and by the recognition that the group process might provide some benefits which individual counseling does not. For example, Rossberg and Jaques (1961) felt that the one-to-one relationship was artificial and insufficient to achieve counseling goals aimed at studying a patient's multiple relationships to the various social groups in his life. Evaluation of group roles in a rehabilitative hospital setting led Rossberg and Jaques to conclude that the impact of a patient's various group relationships has direct significance for his rehabilitative potential.

A policy statement of the American Personnel and Guidance Association (1961), which recommended training and experience in group work as part of a counselor's minimum preparation, gave further indication of the growing awareness of some apparent benefits of group counseling. The same position was reiterated more recently by Johnson (1962) and by Roeber (1962). Wrenn (1960, 1962) indicated that the counselor should see individual and group processes as having different values and stressed the need to distinguish the unique contributions and outcomes of each approach. The following studies reflect some of the more recent experimental efforts in this direction.

Reporting on research being conducted at the University of Illinois, Ohlsen (1960) and Broedel and others (1960) discussed a study in which 29 high-ability ninth graders were randomly assigned to two experimental and two control groups for group counseling and no counseling, respectively. The experimental groups met twice a week for eight weeks. An evaluation of the measures on criteria of academic performance, self-acceptance, and effectiveness of interpersonal relationships indicated significant growth in the experimental groups over the control groups on at least two of the three criteria. Group counseling given to the two control groups after the first experimental period produced similar improvement in one of these groups. Follow-up measures 18 months after counseling showed that the achieved gains had been maintained.

Attitudinal changes similar to those reported by Ohlsen were observed by Cohn and Sniffen (1962) in a pilot study of group counseling with underachieving and misbehaving seventh grade boys. An experimental group of 8 boys, selected from a potential sample of 14 on the basis of ability, willingness to verbalize feelings, and willingness to attend the group meetings, received group counseling twice a week for 10 weeks. Although unable to support their conclusions by adequate postcounseling criterion measures, Cohn and Sniffen felt that the group counseling experience had been successful in achieving the intended purposes of their pilot project.

Unfortunately, limitations of time prevented Cohn and Sniffen from completing their intended experimental design, and this, coupled with their small sample and some apparent bias in selecting the experimental group, raises a question about the six boys who were not selected for counseling. One wonders what changes might have occurred among this noncounseled group during the experimental period and what the relative effectiveness of individual and group counseling might be in dealing with this type of reticent problem student, for whom the prognosis is probably less favorable.

Spielberger, Weitz, and Denny (1962) provided partial answers to these questions about less-anxious students in an elaborate study designed to evaluate the effectiveness of group counseling for improving the academic performance of anxious college freshmen. Incorporating a number of interesting and well-designed experimental controls into their study, the authors demonstrated that the experimental subjects who attended group counseling sessions regularly made significant improvement in academic performance over experimental subjects who did not attend regularly as well as over the noncounseled control group. However, noting that only one-third of those who volunteered for the counseling attended regularly enough to realize any benefits, the authors questioned the applicability of group counseling as a preventive approach to academic underachievement of anxious freshmen. They felt, instead, that group counseling would be indicated primarily for students with personality traits that would enable them to persist in counseling a sufficient number of times to benefit from the experience.

A unique approach to group counseling was presented by Volsky and Hewer (1960). Designed for short-term educational-vocational counseling, the group procedure consisted, essentially, of a case conference approach, in which the educational-vocational plans of each group member were discussed by the group. The goal of each 50-minute session required that some acceptable decision be reached for an appropriate course of action to be followed by the individual discussed. Volsky and Hewer indicated that, although conclusive findings must await further research, experience with over 200 participants had been favorable.

Two studies compared group counseling with individual counseling. One, by Baymur and Patterson (1960), which is reviewed in Chapter IX of this publication, will not be discussed here. The other study, by Stockey (1961), evaluated the relative effectiveness of individual counseling, group counseling, and employment for reducing aggressive behavior among adolescent boys who were experiencing difficulties in adjusting to school. Precounseling and postcounseling measures were used to evaluate changes in adjustment as reflected by attitudes toward self, teacher evaluations, and academic achievement. The three experimental groups were also compared with a control group, with group size ranging from 14 to 16 subjects. Significant improvement in adjustment was shown by subjects of both counseled groups, but neither the employed group nor the control group demonstrated significant change. Group counseling was considered as effective as, but not superior to, individual counseling.

Evaluations of group counseling in conjunction *with* individual counseling were even more sparse than the comparisons *between* the two approaches. While one study by Ford (1962) reported on the use of group and individual counseling as complementary processes with the same students, the procedures outlined in Ford's study seemed to describe a group orientation or group guidance approach when viewed in terms of the process levels suggested by Goldman (1962) or in terms of the connotations of group counseling as found in the other studies reviewed.

Additional References: Gazda and Ohlsen (1961); Waldkoetter (1962).

Applications of Group Dynamics

The extent to which research in group dynamics should be included in this chapter is a moot point, since the October 1959 REVIEW dealt exclusively with that topic. The relations between theory and research in group dynamics and theory and practice in group counseling and guidance, likewise, are not clearly defined. Since recent literature indicates a growing application of group dynamics principles to educational group processes, brief reference to a few pertinent sources seems appropriate.

Two publications of the National Training Laboratories (Bradford, 1961 a, b) pointed to the necessity for applying studies of group dynamics to various psychological, sociological, and educational settings. Bradford (1961a) focused on the general nature of group dynamics and on the use of these forces to help groups function most productively. Bradford (1961b) treated group dynamics in terms of teacher-pupil and other classroom-learning variables. Two additional works, by Lifton (1961) and by Glanz (1962), represented significant treatments of group counseling and guidance functions in terms of group dynamics principles.

Pruitt (1960) summarized the research findings relative to group size and organization and suggested possible implications for the kinds of groups with which the personnel worker is most generally concerned.

Counseling with Parent Groups

Recent literature related to group work with parents changed in focus from that reviewed by Hoyt and Moore in the April 1960 REVIEW. Research during the past three years dealt primarily with programs aimed at helping parents to understand and to utilize test data and other guidance information about their children. In general, reports of such programs indicated favorable responses from the parents who participated, but gave little or no evidence of beneficial carry-over to the respective students.

McCauley (1962) outlined a procedure used for reporting the results of the school's testing program to parents of fourth graders. Parents who attended evening meetings were given the profile sheet of their child's test results. These were explained in a group session. Profiles of underachievers were specially marked so that parents of these children could be encouraged to make an appointment for individual consultation. McCauley did not indicate the percent of parents who responded in any given year or discuss the objectives or any observed values of the program.

Herman and Zeigler (1961) conducted group test interpretations for parents of potential college freshmen. Parents were asked to rate their child on each test area and to predict his first semester grade average. Analysis of parent ratings made before and after the test-interpretation lectures showed some improvement in accuracy of postlecture evaluations. No attempt was made to determine the benefits, if any, that the prospective students might have realized as a result of their parents' involvement. The reported benefits of the program were noted primarily in the favorable opinions expressed by a majority of the participating parents.

Orientation and Special Guidance Courses

Orientation activities probably constituted the most widely used group procedure of the typical guidance functions, but evaluative research with this procedure was, nevertheless, significantly lacking. Reporting on the increased application of group approaches to student personnel services in higher education, Kirkbride (1961) found orientation activities to be among the most relied upon of the group methods reported. Three studies illustrate varied applications of the group orientation procedure.

Duncan (1961) studied the value of group orientation for preparing male freshmen to receive individual counseling. An experimental group of 19 students received two precounseling orientation sessions, while a control group of equal number received only counseling. Postcounseling comparisons between experimental and control subjects showed no differences in terms of self-ratings, realism of vocational goals, and retention of information received in counseling. Although group orientation enabled the counselor to reach more students in a similar time period, it did not increase the effectiveness of the subsequent counseling.

Siegel (1960) described a college placement program that utilized a half-hour group orientation session for disseminating general information to students seeking employment. Since the report did not mention any efforts to evaluate the orientation sessions experimentally, one may assume that the sessions were used primarily for purposes of expediency.

Equally inconclusive in terms of any measured outcome was a program reported by Seymour and Guthrie (1962), in which group orientation was used as a preparation for precollege testing and for test-interpretation interviews. The program was initiated to handle increasing enrollments prior to registration time and to provide practicum subjects for graduate trainees in counselor education. The authors were optimistic about the potential of the program for achieving both objectives.

Glanz, Hayes, and Penney (1959) questioned the adequacy of college orientation programs and proposed a required psychology course as a more appropriate approach to student personnel goals with entering freshmen. Unfortunately, neither type of program received adequate research attention in the literature that has been reviewed. The paucity of supportive evidence, therefore, leaves the significance of orientation programs and of special guidance courses dependent upon more stringent evaluation.

Occupation Units, Career Conferences, and Vocational Guidance in Regular Course Work

Many group guidance functions traditionally have been carried on as special units of regular subject matter courses, but research reports of their efficacy have long been conflicting and inadequate.

One of the best designed studies on this topic in recent literature was by Rusalem and Darer (1960). The study, which represented considerably more than the typical subject matter unit, was conducted in conjunction with a double-period, modified English class over a two-year period. During this time, the experimental classes spent one 80-minute period each week discussing topics related to vocational guidance goals. In addition to the modified curriculum, the experimental groups received individual and group counseling, teacher consultation, and parent counseling. Six class groups of tenth graders, with 25 boys and girls in each group, constituted the study sample. Three of the groups made up the experimental population, and the other three served as controls. Comparisons on several criteria showed significant differences between experimental and control groups, with all gains being in favor of the experimental group.

Occupation courses and career days were apparently very low on the hierarchy of research interests. Although surveys by Sinick and Hoppock (1960) and by Carter and Hoppock (1961) indicated that occupation courses were being offered regularly in many states and at a number of universities, there was a dearth of research evaluating this particular guidance offering during the past three years. Equally sparse were studies dealing with the subject of career days. However, Shulim (1959) discussed an experiment with career-curriculum conferences at the college level and outlined a number of modifications that were developed through "post-mortem examinations" of the program after each of three experimental trials.

In view of the many known secondary schools that incorporate annual career days and occupation units as regular guidance activities, more research information regarding these procedures is greatly needed. Additional applications at the college level also warrant further investigation. Sinick and Hoppock (1960) encouraged research in the teaching of occupations and offered to summarize and to publish reports of any such research.

While not reporting on specific research, Caldwell (1960) and Mahler and Caldwell (1961) drew upon their experiences with a number of experimental programs to help teachers and counselors to implement group guidance and group counseling procedures within daily classroom activities.

Additional References: Brundage and Frank (1961); Speary (1959).

Summary and Implications

The mounting research evidence in support of group counseling cannot be ignored. However, the distinctive merits of group counseling as compared with individual counseling are still inconclusive. While several studies outlined some of the benefits that their authors believed group counseling could provide, the assumed benefits of the group experience were not often tested as part of the experimental design. What remains to be demonstrated more conclusively is the nature of situations and conditions for which multiple as well as individual counseling is most appropriately indicated. The values of individual and group counseling as concurrent processes also need further investigation.

The influence of group dynamics literature and the importance given to the relationship in counseling raises several other questions relative to needed research: What kinds of relationships are established in group counseling, and how can these be identified and measured? How are individual and group roles influenced by the relationships established? How do various roles affect the outcomes of the group process? Studies by Altman and McGinnies (1960), by Cohn, Ohlsen, and Proff (1960), by Noble, Ohlsen, and Proff (1961), and by Ohlsen and Oelke (1962) give some direction for seeking research answers to these questions.

Several studies suggested the need to identify criteria which are more sensitive to changes occurring among normal populations—changes that might be expected after one or two guidance contacts as contrasted with criterion changes anticipated after longer-term, more intensive relationships. Difficulty in measuring slight and more immediate changes in the behavior of normal students may well account for the lack of significant research with group guidance procedures as well as the descriptive rather than experimental nature of the available studies.

Finally, it seems an unfortunate commentary on the research sophistication of school counselors to note that an overwhelming preponderance of the research literature in guidance and counseling is still being reported.

if not undertaken, by counselor-educators rather than by pupil personnel workers. Inasmuch as group guidance procedures have special relevance for elementary and secondary schools, the importance of more significant research from school practitioners cannot be overemphasized.

Bibliography

ALTMAN, IRWIN, and MCGINNIES, ELLIOTT. "Interpersonal Perception and Communication in Discussion Groups of Varied Attitudinal Composition." *Journal of Abnormal and Social Psychology* 60: 390-95; May 1960.

AMERICAN PERSONNEL AND GUIDANCE ASSOCIATION. "Standards for the Preparation of School Counselors." *Personnel and Guidance Journal* 40: 402-407; December 1961.

BAYMUR, FERIHA B., and PATTERSON, CECIL H. "A Comparison of Three Methods of Assisting Underachieving High School Students." (Comment by David Tiedeman.) *Journal of Counseling Psychology* 7: 83-90; Summer 1960.

BRADFORD, LELAND P., editor. *Group Development*. Selected Reading Series One. Washington, D.C.: National Training Laboratories, National Education Association, 1961. 106 pp. (a)

BRADFORD, LELAND P., editor. *Human Forces in Teaching and Learning*. Selected Reading Series Three. Washington, D.C.: National Training Laboratories, National Education Association, 1961. 102 pp. (b)

BROEDEL, JOHN, and OTHERS. "The Effects of Group Counseling on Gifted Underachieving Adolescents." *Journal of Counseling Psychology* 7: 163-70; Fall 1960.

CALDWELL, EDSON. *Group Techniques for the Classroom Teacher*. Chicago: Science Research Associates, 1960. 86 pp.

CARTER, EDWARD M., and HOPPOCK, ROBERT. "College Courses in Careers." *Personnel and Guidance Journal* 39: 373-75; January 1961.

COHN, BENJAMIN, and SNIFFEN, A. MEAD. "A School Report on Group Counseling." *Personnel and Guidance Journal* 41: 133-38; October 1962.

COHN, BENJAMIN; OHLSEN, MERLE M.; and PROFF, FRED. "Roles Played by Adolescents in an Unproductive Counseling Group." *Personnel and Guidance Journal* 38: 724-31; May 1960.

DAVIS, DONALD A. "Effect of Group Guidance and Individual Counseling on Citizenship Behavior." *Personnel and Guidance Journal* 38: 142-45; October 1959.

DUNCAN, HELEN GENE. *The Effect of Group Orientation on the Outcomes of Vocational Counseling*. Doctor's thesis. Gainesville: University of Florida, 1961. 205 pp. Abstract: *Dissertation Abstracts* 22: 487; No. 2, 1961.

FORD, DONALD H. "Group and Individual Counseling in Modifying Behavior." *Personnel and Guidance Journal* 40: 770-73; May 1962.

GLANZ, EDWARD C. *Groups in Guidance: The Dynamics of Groups and the Application of Groups in Guidance*. Boston: Allyn and Bacon, 1962. 385 pp.

GLANZ, EDWARD C.; HAYES, ROBERT W.; and PENNEY, JAMES F. "The Freshman Psychology Course as the Basis for a Student Personnel Program." *Personnel and Guidance Journal* 38: 290-95; December 1959.

GOLDMAN, LEO. "Group Guidance: Content and Process." *Personnel and Guidance Journal* 40: 518-22; February 1962.

HERMAN, LOUIS M., and ZEIGLER, MARTIN L. "The Effectiveness of Interpreting Freshman Counseling-Test Scores to Parents in a Group Situation." *Personnel and Guidance Journal* 40: 143-49; October 1961.

JOHNSON, WALTER F. "Position Paper—Curriculum." *Counselor Education: A Progress Report on Standards*. Washington, D.C.: American Personnel and Guidance Association, 1962. pp. 13-18.

KEMP, GEORGE ALLEN. *A Comparative Evaluation of the Effectiveness of Group Guidance and Individual Counseling with Freshmen. Doctor's thesis*. Denton: North Texas State University, 1962. 150 pp. Abstract: *Dissertation Abstracts* 22: 4270; No. 12. 1962.

KIRKBRIDE, VIRGINIA R. "Group Approaches to Student Personnel Services in Higher Education." *Journal of the National Association of Women Deans and Counselors* 24: 127-31; April 1961.

Lifton, Walter M. *Working with Groups: Group Process and Individual Growth.* New York: John Wiley & Sons, 1961. 238 pp.

McCauley, John H. "Reporting Results of the Standardized Testing Program to Parents." *Personnel and Guidance Journal* 41: 56-57; September 1962.

Mahler, Clarence A., and Caldwell, Edson. *Group Counseling in Secondary Schools.* Chicago: Science Research Associates, 1961. 78 pp.

Noble, Frank; Ohlsen, Merle M.; and Proff, Fred. "A Method for the Quantification of Psychotherapeutic Interaction in Counseling Groups." (Comment by Leon Gorlow.) *Journal of Counseling Psychology* 8: 54-61; Spring 1961.

Ohlsen, Merle M. "Counseling Within a Group Setting." *Journal of the National Association of Women Deans and Counselors* 23: 104-109; April 1960.

Ohlsen, Merle M., and Oelke, Merritt C. "An Evaluation of Discussion Topics in Group Counseling." *Journal of Clinical Psychology* 18: 317-22; July 1962.

Pruitt, Wilton. "Group Size and Organizational Planning." *Personnel and Guidance Journal* 38: 626-32; April 1960.

Rinn, John L. "Group Guidance: Two Processes." *Personnel and Guidance Journal* 39: 591-94; March 1961.

Roeber, Edward C. "Position Paper—Practicum and Internship." *Counselor Education: A Progress Report on Standards.* Washington, D.C.: American Personnel and Guidance Association, 1962. pp. 24-31.

Rossberg, Robert, and Jaques, Marceline. "The Role of the Group in Patient Evaluation, Counseling, and Management." *Personnel and Guidance Journal* 40: 135-42; October 1961.

Rusalem, Herbert, and Darer, Hyman. "Vocational Counseling of the Slow Learner." *Journal of the National Association of Women Deans and Counselors* 23: 110-16; April 1960.

Seymour, John C., and Guthrie, Fain A. "A Summer Program Combining Freshman Orientation, Counseling, and Counselor Preparation." *Personnel and Guidance Journal* 40: 477-79; January 1962.

Shulim, Joseph I. "Experimenting with a Career-Curricular Conference." *Personnel and Guidance Journal* 38: 222-25; November 1959.

Siegel, Max. "Group Orientation and Placement Counseling." *Personnel and Guidance Journal* 38: 659-60; April 1960.

Sinick, Daniel, and Hoppock, Robert. "Research by States on the Teaching of Occupations." *Personnel and Guidance Journal* 39: 218-19; November 1960.

Spielberger, Charles D.; Weitz, Henry; and Denny, J. Peter. "Group Counseling and the Academic Performance of Anxious College Freshmen." *Journal of Counseling Psychology* 9: 195-204; Fall 1962.

Stockey, Merrel Richard. *A Comparison of the Effectiveness of Group-Counseling, Individual-Counseling, and Employment Among Adolescent Boys with Adjustment Problems.* Doctor's thesis. Ann Arbor: University of Michigan, 1961. 97 pp. Abstract: *Dissertation Abstracts* 22: 491; No. 2, 1961.

Volsky, Theodore, and Hewer, Vivian H. "A Program of Group Counseling." *Journal of Counseling Psychology* 7: 71-73; Spring 1960.

Wrenn, C. Gilbert. "Editorial Comment." *Journal of Counseling Psychology* 7: 242; Winter 1960.

Wrenn, C. Gilbert. *The Counselor in a Changing World.* Washington, D.C.: American Personnel and Guidance Association, 1962. 195 pp.

Additional References

Brundage, Erven, and Frank, Stanley D. "The Career Conference Concept in San Diego County." *Personnel and Guidance Journal* 40: 174-76; October 1961.

Gazda, George, and Ohlsen, Merle. "The Effects of Short-Term Group Counseling on Prospective Counselors." *Personnel and Guidance Journal* 39: 634-38; April 1961.

Speary, William. "Good Guidance Through Career Days." *American Vocational Journal* 34: 21-22; December 1959.

Waldkoetter, Raymond O. "A Comparative Guidance Study: Group Counseling with Selected African Student-Teachers." *Personnel and Guidance Journal* 40: 638-42; March 1962.

CHAPTER IX

Program Evaluation

C. H. PATTERSON *

THE IMPORTANCE of and the necessity for evaluation of guidance and counseling programs and services have been repeatedly stressed. But as Rothney and Farwell noted in the April 1960 REVIEW, there is little evidence that much is being done in this area. In the April 1957 REVIEW, Cottle also commented on the paucity of research relating to evaluation. The present chapter will indicate the continuing scarcity of studies evaluating counseling and guidance services. However, although such studies continue to be few in number, the quality seems to be improving, with relatively more of them being experimental investigations rather than survey or simple follow-up studies.

For the purposes of this chapter, guidance services are defined as personnel services related to and including counseling, performed by a professional staff apart from the instructional program but in a school setting. The writer is not concerned with other personnel services, such as school psychology, or with psychotherapy. Furthermore, since the chapter is limited to studies evaluating the outcomes of services, it does not include the studies of processes or procedures that are covered in Chapters V and VIII.

The literature will be considered in the following organization: general discussions, follow-up studies, control group studies, methodological problems, and summary and conclusions.

General Discussions

Continuing a series of publications, Hill and Nitzschke (1960) presented procedures for obtaining evaluations of a guidance program from students and parents and provided a model study. Hill (1960) furnished a similar study of a school's testing program by its staff. The Oklahoma State Department of Education (1961) included a chapter on evaluation in its handbook. Wellman and Twiford (1961) offered suggestions for state and local evaluative studies to meet the requirements of Title V-A of the National Defense Education Act of 1958. Endicott (1961) provided a check list of evaluative criteria for a guidance program for superior students.

---

* Appreciation is expressed to Marilyn Dunn for assistance in searching for and reviewing the literature.

Follow-Up Studies

Various classifications of evaluative studies have been proposed. For the purposes of this chapter, a classification into follow-up or ex post facto studies and experimental or controlled studies is adequate.

Junior and Senior High School Studies

There were few studies of the outcomes of guidance activities by means of simple follow-up. Moore (1960) administered a test of guidance information to freshmen, and again to the same students as seniors, and found a significant gain in knowledge. Students who had more than one counselor gained more; and, surprisingly, those who had untrained counselors gained more than those who had counselors with some training. Gribbons (1960) reported on the use of the manual and guide, *You: Today and Tomorrow*, in nine eighth grade classes taught by seven teachers. Fifty-five of the boys and 53 of the girls were interviewed before and after use of the material and rated for various factors pertinent to vocational choice. All scores were significantly higher at the second interview. Rubinfeld and Hoppock (1961) obtained opinions of a ninth grade course from students in occupations eight years later; although ratings of the course were not highly favorable, 80 percent felt that it should be a required course. Carlin (1960) found that only 40 percent of a group of college freshmen had had a course on the world of work while in high school, and less than 50 percent of these reported that it had been helpful.

Additional References: Cohn and Sniffen (1962); Impellizzeri (1960); Thomson (1960); Wrightstone (1960).

College Studies

Glazer (1960) obtained questionnaire responses from 118 of 500 former clients of a university counseling center at least 18 months following their experience; 90 percent declared that they were satisfied or very satisfied. One wonders whether the same proportion of the nonrespondents felt the same way. Apostal (1960) obtained occupational data from 114 students (81-percent return) who had had vocational counseling. Three judges rated appropriateness of occupational choice and adequacy of the counseling process on the basis of the counseling records. Eighty percent of the occupational choices were rated as appropriate, and the counseling process was rated as adequate in 67 percent of the cases. Without a control group, ratings of appropriateness of vocational choice obviously cannot be taken as evidence of counseling adequacy. Gonyea (1962) compared precounseling and postcounseling vocational objectives of 227 clients of a university counseling center with their objectives stated one year after counseling. Nearly one-half changed their vocational plans between precounseling and postcounseling. About 60 percent of these changes were in the direction of greater appropriateness, as noted by four judges. One year later, only 29 percent were still following their postcounseling objectives. However, those students whose postcounseling objectives were rated appropriate had objectives one year later that were more frequently rated as appropriate. There was no improvement in grades following counseling, either for the total sample or for those whose vocational choices became more appropriate. Gonyea concluded that appropriateness of vocational choice does not appear to be a valid measure of enduring, meaningful effects of vocational counseling.

Additional Reference: Hackett (1960).

Control Group Studies

Junior and Senior High School Studies

Shimberg and Katz (1962) reported the results of the use of the manual and guide, *You: Today and Tomorrow*, with eighth and ninth grade students. Scores on a guidance test following use of the material in the classroom were significantly higher than the scores of an equated group which did not participate in the program—a finding indicating that the program results in an acquisition of information or knowledge related to guidance concepts. Davis (1959) divided 30 students low in citizenship grades into three random groups. Group A had twenty 50-minute periods of group guidance; Group B had two periods each of individual counseling; and Group C constituted the control group. Counseling was nondirective. Group A showed significantly greater improvement in citizenship grades than did Group B or C, although Group B also gained significantly. While the study might appear to show superiority for the group approach, it should be noted that while Groups A and B were equated in terms of total counselor time, individuals in Group A had 10 times as much counseling time. In another study, Davis (1962) compared two groups of 21 students each who were potential dropouts. One group was given special attention, which included special counseling sessions. Members of the experimental group received fewer failing grades. They were referred to the principal less frequently and for less serious problems. Moreover, there were fewer dropouts in the experimental group. It is, of course, impossible to know to which aspects of the treatment the results may be attributed.

Baymur and Patterson (1960) compared four groups of underachievers, eight students in each group. One sample had group counseling (8 sessions); one, individual counseling (9-12 sessions); one, a single motivational counseling session; and one served as a control group. Although there were no significant differences among the four groups, the two counseled groups combined showed an increase in Q-sort adjustment score and in grade average. Broedel and others (1960) reported a study of group counseling of gifted underachievers based on Broedel's thesis, which has been reviewed by Hoyt and Moore in the April 1960 REVIEW. Following 16 group counseling sessions, there was increased acceptance of self and others, improved scores on an achievement test, and improved ability to relate to peers, siblings, and parents; but there was no improvement in grades.

Arbuckle and Boy (1961) compared three groups of students classified as behavior problems. One group received client-centered group counseling; the second, after-school detention; and the third, neither experience. Although the counseling was not voluntary, it appeared to be effective.

Additional Reference: Lurie, Goldfein, and Baxt (1960).

College Studies

Richardson (1960) found no differences in grade point average, number of dropouts, or number of transfers, in two matched groups of engineering students in the lowest one-third of their class. One group received an average of 2.5 information- and advice-giving interviews; the other groups, no special attention. Goodstein and Crites (1961) also found no apparent effect of two to five educational-vocational counseling interviews upon grade point averages of students in a probationary summer session. Moore and Popham (1960) compared a group given three student-centered interviews, a group given three content-centered interviews, and a control group given no interviews. There were no differences in grade point averages, but the group having the student-centered interviews gained significantly in average score on the inventory of academic adjustment given at the beginning and the end of the semester.

Williams (1962) provided brief exposure of educational-vocational counseling, consisting of two interviews and a testing session, to a group of applicants and compared them on adjustment measures with (a) a control group of applicants who were not given counseling and (b) a control group who had not sought counseling. Counseling led to an increase in Q-sort adjustment scores and in congruence among concepts of self, ideal self, and ordinary person, which persisted four months after counseling.

Watson (1961) compared equivalent groups of students with personal and emotional problems who (a) accepted counseling; (b) appeared for one interview, but denied the need for help; and (c) did not respond to the invitation to appear for an interview. After two years, the first group was superior to the other groups in scholastic achievement and in other respects indicated by college records. Searles (1962) compared two groups of superior college freshmen, one of which was given three interviews during the semester, in which they were informed of their selection for the study because of their high ability and in which their educational-vocational objectives were discussed. There were no differences between the groups in grade point average, number completing the semester, number placed on academic probation, or number placed on the dean's list for scholarship. Stamatakos and Shaffer (1959) compared four groups of superior female freshmen students; each group received different degrees of special attention. One of them (an experimental group) received monthly structured personal counseling interviews with dormitory counselors. The groups did not differ significantly in grade point average or in participation in activities, although the students in the experimental group rated the program highly.

In a well-designed and well-executed study, Spielberger, Weitz, and Denny (1962) found that anxious college freshmen who volunteered for and participated in group counseling (between 2 and 11 sessions per student) improved significantly more in grade point average than did a control group who volunteered for but did not receive counseling.

Methodological Problems

The recommended procedures for evaluative studies, such as those found in Patterson (1962) and in Wellman and Twiford (1961), may be outlined as follows:

1. Select and describe the objectives or outcomes that the program or service should achieve.
2. Define the criteria that are acceptable as indicators of the achievement or lack of achievement of these objectives.
3. Select or devise instruments to measure the defined criteria.
4. Design the experiment in terms of the samples and their selection, the experimental or statistical controls, the administration or identification of the treatment or independent variable, and the application of the criterion measures.
5. Perform a statistical analysis of the data to determine the effectiveness of the treatments in terms of the outcome criteria.

Difficulties are encountered at several points in attempting to implement this procedure.

Criterion Problem

The objectives or goals of counseling and guidance services are difficult to specify. The problem is not so simple as Tyler (1961) implies in her "question of whether the criterion of counseling success should be satisfaction or adjustment" (p. 261). In the first place, there are more choices; and in the second place, it is questionable that either satisfaction or adjustment is an acceptable goal (Patterson, 1958; Goodstein and Grigg, 1959).

The desire for specific and measurable criteria has led to the proliferation of easily observable and objective outcomes. These include such factors as reduction in dropouts, increase in high school graduates going to college, improvement in grades, reduction in underachievement, selection of a vocational objective, stability of vocational choice, stability in employment, salary level, and salary increase. However, objections may be and have been raised to all of these criteria (e. g., Miller, 1961). Not all students should remain in school until graduation; not all graduates should go to college; not all students should select a vocational objective before high school graduation. Change in vocational choice may be desirable. Underachievers may have (a) values and goals other than academic and (b) certain talents that should be developed.

Wolfle (1960) has made the point that everyone should be aided in attaining the highest level he can reach in the area in which he has the greatest talent and interest. In some instances, striving for academic achievement or even trying to be a well-rounded individual may be a secondary goal. Thus, if a student tends to neglect study of academic subjects in order to practice the piano and thus becomes an outstanding pianist, he may contribute more to society and to his own mental health and satisfaction than if he attempts to excel in academic subjects. In addition, grade point average is an inappropriate criterion for counseling directed to achieve better personal adjustment. Moreover, personal adjustment may actually suffer following such counseling.

The point is that none of these criteria can be accepted as adequate for all individuals. Investigators must move on to the measurement of effects relevant to the individual. Indeed, the development of the potentialities of the individual through self-actualization seems to be generally accepted as the objective of counseling and guidance services, though it is often lost in the desire to use a tangible, objective criterion. Inherent in this objective is the development of an adequate self-concept. This includes awareness and understanding of one's self and attitudes, the development of attitudes of self-esteem and self-respect, the achievement of independence and self-direction, and the ability to take responsibility for oneself and for one's decisions and actions. Progress has been made in measuring such factors, although the profession does not have adequate instruments, in terms of reliability and demonstrated validity. However, it would appear to be more profitable to struggle with the task of converting such concepts and goals into measuring instruments than to continue employing the objective but sterile criteria which are used too often. There is evidence that researchers are moving in this direction. Attempts to measure the self-concept by such instruments as Q-sorts have been reported, for example, in a study by Sheldon and Sorenson (1960).

Miller (1961) suggested a different approach from the generally accepted one, in order to avoid the problem of determining or agreeing upon objectives with their value connotations of goodness or badness. In this empirical approach, specific guidance procedures or services would be provided, and then differences between groups receiving services and those not receiving services would be sought. This, of course, is actually a commonly used approach. Although it might avoid the problem of agreeing upon desirable goals, it does not avoid the problem of determining what to measure and how to measure it.

Problems of Experimental Design

Research in counseling and guidance encounters the basic problems of experimentation with human beings. Patterson (1960) continued an earlier discussion (Patterson, 1956) of these problems. In addition to the difficulties posed by sampling, there are the problems of controls. It is not usually acceptable to withhold a service from one group in order to have a basis for determining its effects on another group, although it might be argued that this deprivation would not be great. However, it may not be necessary to withhold services, since other methods, such as delayed treatment, the own-control method, and the provision of differential treatment to comparable groups, are useful (Patterson, 1960). However, the problem of equating groups remains. Randomization of sampling or of the application of treatments and use of matching techniques pose difficulties (Patterson, 1956). All relevant variables, such as motivation, changes in family pressures and finances, and other environmental changes, usually cannot be controlled or measured. They may operate to negate or to affect differentially the effects of treatments. As a result, it is seldom possible to be sure that lack of results really means that the treatment was ineffective, or that results obtained are really due to the treatments. Nevertheless, it is encouraging to see, in the studies reviewed, the increasing concern with attempts to achieve control of unwanted sources of variance.

A factor that has not been adequately recognized is the nature of the treatment variable, that is, the kinds of services provided. One aspect of this problem is the impossibility of being able to determine what aspect of a program is responsible for any obtained results when the treatment consists of a complex of services or activities (Patterson, 1955). This confounding was present in a number of studies reviewed (Apostal, 1960; Davis, 1962; Glazer, 1960; Rubinfeld and Hoppock, 1961; Stamatakos and Shaffer, 1959).

A second aspect of this problem is present even when only one service, such as counseling, is provided. Counseling is not always counseling—there are differences, not only among individuals but also among general approaches or points of view. One must specify the nature of the counseling involved; few studies do this. Too many studies do not indicate the training and experience of the counselors involved. When such information has been indicated, it appears that most of the research has involved partially trained and inexperienced counselors.

Tyler (1961) stated that no additional investigations are needed in which directive and nondirective approaches are compared, since a few studies have shown no differences. However, these studies are all questionable because of the use of inexperienced counselors and counselors employing an approach in which they had no confidence or real understanding. Moreover, having the same counselor use both approaches, which ostensibly controls the counselor personality variable, actually introduces error related to differences in counselor attitudes, experience, and skill in different approaches. In most studies one does not know what the treatment has been when counseling is said to have been provided.

Summary and Conclusions

Although progress in the evaluation of counseling and guidance services is present, it has been relatively limited during the past three years. There has been an increase in controlled studies, but insufficient attention has been paid to methodological aspects of such research. Contradictory results are prevalent. These are probably related to differences in samples as well as to other factors noted in this chapter, such as duration and intensity of the services provided, nature or type of the services provided, and experience and skill of the counselor.

As in the past, some studies have found counseling and guidance services to be effective. In other studies, the results did not negate the null hypothesis. Some of the best controlled studies have yielded minimal or negative results. This state of affairs is not necessarily to be taken as support for the null hypothesis that counseling and guidance services make no difference. Rather it is related to a number of factors that make it impossible to compare studies or draw any conclusions from them as a group. These factors include the use of different and sometimes inappropriate criteria, some of which are inconsistent with the goals and objectives of education and counseling. In some cases, the criteria chosen are unreasonable in terms of what might be expected from the services provided. Berdie (1956) has pointed out "expected changes in larger goal-related behaviors may be naïvely optimistic or presumptuous" (p. 346) following short-term counseling. It is hard to believe that investigators expect results from one or two interviews. Such studies are usually negative (e. g., Moore and Popham, 1960; Richardson, 1960; Searles, 1962) except when the criterion is the acquisition of simple information or when vocational counseling of normal individuals is involved (Williams, 1962). Some of the best controlled studies that have reported minimum results (e. g., Broedel and others, 1960; Baymur and Patterson, 1960) have suggested that the extent or duration of the counseling involved was too limited.

As mentioned previously, another factor contributing to lack of results has been the use of inexperienced and inadequately trained counselors. The finding that such counselors are ineffective has no necessary relation to what might be found when experienced and well-trained counselors are used.

From the analysis of the deficiencies of current research, a number of suggestions or recommendations for future research are apparent:

1. Consideration must be given to the goals and objectives of counseling and guidance services and to the criteria relevant to the attainment of these goals and objectives developed and used in future studies.
2. Attention must be given to specifying and to defining the nature of the treatment variable in order that studies may be replicated and in order that one may know to what variables any results may be attributed. Study of specific, defined methods or services in terms of specific criteria will lead to knowledge of what leads to what and will enable investigators to select methods or approaches that will lead to desired criteria or outcomes (e. g., Davis, 1959).
3. An adequate test of the effects of counseling, especially when criteria of personality changes are used, must provide counseling services that are sufficiently extensive and intensive to provide realistic expectations for such changes. It is unreasonable to expect superficial one-interview counseling to have such effects.
4. Any adequate test of the influences or effects of counseling must be based upon the use of counselors who are trained and experienced and who have confidence in the methods or approaches they use.
5. Although it is of interest to study the effects of counseling on unmotivated clients or on clients who do not apply or volunteer for counseling, the primary concern is with individuals who are interested in or desirous of receiving counseling. Studies using involuntary clients are not a test of the effects of counseling in a normal counseling situation.
6. Long-term follow-up is necessary to ascertain the nature and persistence of effects. In some instances, there are delayed effects; in others, there may be superficial effects immediately following counseling which will not persist.

Controlled experiments which meet these requirements are difficult and expensive to conduct, not only in terms of experimental design and controls, but also in terms of time, including the duration of the experiment and the follow-up period. It would appear that, as in other areas of research, an adequate study requires more than the resources of a single investigator. The time is ripe for an extensive, long-term investigation with adequate financial spport, in which existing knowledge may be applied to the conduct of meaningful research on the effects of counseling and guidance services.

Bibliography

APOSTAL, ROBERT A. "Two Methods of Evaluating Vocational Counseling." *Journal of Counseling Psychology* 7: 171-75; Fall 1960.

Arbuckle, Dugald S., and Boy, Angelo V. "Client-Centered Therapy in Counseling Students with Behavior Problems." *Journal of Counseling Psychology* 8: 136-39; Summer 1961.

Baymur, Feriha B., and Patterson, Cecil H. "A Comparison of Three Methods of Assisting Underachieving High School Students." (Comment by David V. Tiedeman.) *Journal of Counseling Psychology* 7: 83-90; Summer 1960.

Berdie, Ralph F. "Counseling." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth.) Palo Alto, Calif.: Annual Reviews, 1959. Vol. 10, pp. 345-70.

Broedel, John, and others. "The Effects of Group Counseling on Gifted Underachieving Adolescents." *Journal of Counseling Psychology* 7: 163-70; Fall 1960.

Carlin, Leslie O. "Vocational Decisions and High School Experiences." *Vocational Guidance Quarterly* 8: 168-70; Spring 1960.

Davis, Donald A. "Effect of Group Guidance and Individual Counseling on Citizenship Behavior." *Personnel and Guidance Journal* 38: 142-45; October 1959.

Davis, Donald A. "An Experimental Study of Potential Drop-Outs." *Personnel and Guidance Journal* 40: 799-802; May 1962.

Endicott, Frank Simpson. *Guiding Superior and Talented High School Students.* Superior and Talented Student Project, North Central Association. Minneapolis: Lund Press, 1961. 84 pp.

Glazer, Stanford H. "Counselee Attitude Toward Counseling." *Vocational Guidance Quarterly* 8: 235-38; Summer 1960.

Gonyea, George C. "Appropriateness-of-Vocational-Choice as a Criterion of Counseling Outcome." *Journal of Counseling Psychology* 9: 213-19; Fall 1962.

Goodstein, Leonard D., and Crites, John O. "Brief Counseling with Poor College Risks." *Journal of Counseling Psychology* 8: 318-21; Winter 1961.

Goodstein, Leonard D., and Grigg, Austin E. "Client Satisfaction, Counselors, and the Counseling Process." (Comment by C. H. Patterson.) *Personnel and Guidance Journal* 38: 19-24; September 1959.

Gribbons, Warren D. "Evaluation of an Eighth Grade Group Guidance Program." *Personnel and Guidance Journal* 38: 740-45; May 1960.

Hill, George E. *The Staff Evaluate the School's Testing Program.* Athens: Center for Educational Service, Ohio University, 1960. 30 pp.

Hill, George E., and Nitzschke, Dale F. *Students and Parents Evaluate the School's Guidance Program.* Athens: Center for Educational Service, Ohio University, 1960. 23 pp.

Miller, Carroll H. *Foundations of Guidance.* New York: Harper & Row, 1961. 464 pp.

Moore, Gilbert D. "An Investigation of Certain Aspects of a Guidance Program." *Personnel and Guidance Journal* 38: 558-62; March 1960.

Moore, Mary R., and Popham, W. James. "Effects of Two Interview Techniques on Academic Achievement." *Journal of Counseling Psychology* 7: 176-79; Fall 1960.

Oklahoma State Department of Education. *A Handbook for the Improvement of Guidance and Counseling in Oklahoma Schools: Grades K-12.* Oklahoma City: the Department, 1961. 102 pp.

Patterson, Cecil H. "Comment on *An Evaluation of the Counseling of Chronic Schizophrenics* by Bernard A. Stotsky, Paul G. Daston, and C. Nicholas Vardock." *Journal of Counseling Psychology* 2: 254-55; Winter 1955.

Patterson, Cecil H. "Matching Versus Randomization in Studies of Counseling." *Journal of Counseling Psychology* 3: 262-71; Winter 1956.

Patterson, Cecil H. "Client Expectations and Social Conditioning." *Personnel and Guidance Journal* 37: 136-38; October 1958.

Patterson, Cecil H. "Methodological Problems in Evaluation." *Personnel and Guidance Journal* 39: 270-74; December 1960.

Patterson, Cecil H. *Counseling and Guidance in Schools: A First Course.* New York: Harper & Row, 1962. 382 pp.

Richardson, LaVange H. "Counseling the Ambitious Mediocre Student." *Journal of Counseling Psychology* 7: 265-68; Winter 1960.

Rubinfeld, William A., and Hoppock, Robert. "Occupations Course Evaluated Eight Years Later." *Vocational Guidance Quarterly* 10: 45-48; Autumn 1961.

Searles, Aysel, Jr. "The Effectiveness of Limited Counseling in Improving the Academic Achievement of Superior College Freshmen." *Personnel and Guidance Journal* 40: 630-33; March 1962.

SHELDON, M. STEPHEN, and SORENSON, A. GARTH. "On the Use of Q-Technique in Educational Evaluation and Research." *Journal of Experimental Education* 29: 143-51: December 1960.

SHIMBERG, BENJAMIN, and KATZ, MARTIN R. "Evaluation of a Guidance Text." *Personnel and Guidance Journal* 41: 126-32; October 1962.

SPIELBERGER, CHARLES D.; WEITZ, HENRY; and DENNY, J. PETER. "Group Counseling and the Academic Performance of Anxious College Freshmen." *Journal of Counseling Psychology* 9: 195-204; Fall 1962.

STAMATAKOS, LOUIS C., and SHAFFER, ROBERT H. "Effects of Special Attention upon Potentially Superior Freshman Students." *Personnel and Guidance Journal* 38: 106-11; October 1959.

TYLER, LEONA E. *The Work of the Counselor.* Second edition. New York: Appleton-Century-Crofts, 1961. 327 pp.

WATSON, GLADYS H. "An Evaluation of Counseling with College Students." *Journal of Counseling Psychology* 8: 99-104; Summer 1961.

WELLMAN, FRANK E., and TWIFORD, DON D. *Guidance Counseling and Testing Program Evaluation: Suggestions for Secondary Schools.* U.S. Department of Health, Education, and Welfare, Office of Education. Washington, D.C.: Government Printing Office, 1961. 37 pp.

WILLIAMS, JOHN E. "Changes in Self and Other Perceptions Following Brief Educational-Vocational Counseling." (Comment by William U. Snyder.) *Journal of Counseling Psychology* 9: 18-30; Spring 1962.

WOLFLE, DAEL. "Diversity of Talent." *American Psychologist* 15: 535-45; August 1960.

Additional References

COHN, BENJAMIN, and SNIFFEN, A. MEAD. "A School Report on Group Counseling." *Personnel and Guidance Journal* 41: 133-38; October 1962.

HACKETT, HERBERT R. "Non-Student Cases at a University Counseling Center." *Vocational Guidance Quarterly* 8: 180-83; Spring 1960.

IMPELLIZZERI, IRENE H. "The New York City Talent Preservation Project: An Experiment in Guidance of the Mentally Superior." *Journal of the National Association of Women Deans and Counselors* 23: 99-103; April 1960.

LURIE, WALTER A.; GOLDFEIN, JACOB; and BAXT, ROLAND. "An Intensive Vocational Counseling Program for Slow Learners in High School." *Personnel and Guidance Journal* 39: 21-29; September 1960.

THOMSON, MURIEL. "Students' Evaluations of Guidance Services." *California Journal of Educational Research* 11: 167-69; September 1960.

WRIGHTSTONE, J. WAYNE. "Demonstration Guidance Project in New York City." *Harvard Educational Review* 30: 237-51; Summer 1960.

Index to Volume XXXIII, No. 2

Page citations are made to single pages; these are often the beginning of a chapter, section, or running discussion dealing with a topic.

Appraisal function in relation to: communicating appraisal information, 193; evaluation of test reports, 193; interpreting data, 190; nontest methods, 190; scoring and score conversions, 189; test taking and its preparation, 188

Career conferences: 210

Career development: general, 197; in women, 200

Concepts and purposes underlying guidance and personnel practice: in colleges and universities, 147; in schools, 147

Counseling: avowed goals, 179; diagnosis, 183; methods, 183; obtained results, 180; theory construction, 179

Counseling procedures, group: 205

Elementary and secondary school programs: issues relating to coordination of pupil personnel services, 159; organization and administration of guidance services, 152; roles and functions of pupil personnel workers, 156

Foundations in guidance, counseling, and personnel services: concepts and purposes of, 146; philosophical trends in, 146; society and culture in relation to, 143

Group counseling: 205

Group dynamics: applications of, 208

Group procedures: applications of group dynamics, 208; counseling with parent groups, 208; group counseling, 205; occupations units, career conferences, and vocational guidance in regular course work, 210; orientation and special guidance courses, 209

Higher education programs: in extended education, 166; in four-year colleges, 163; in junior colleges, 166; in relation to issues underlying research needs, 168

Interpreting data in the appraisal function: counselor vs. statistical prediction, 191; differential predictability, 192; expectancy studies, 190; use of *Strong Vocational Interest Blank*, 191

Issues relating to coordination of pupil personnel services: 159

Methodological problems in program evaluation: criterion problem, 218; in relation to experimental design, 220

Occupational classification: 200

Occupational information, sources of: 201

Occupational psychology: 199

Occupations units, career conferences, and vocational guidance in regular course work: 210

Organization and administration of guidance services in relation to: special services, 154; teacher vs. specialist participation, 155; timing in introducing guidance services, 155; types of counselor assignments, 153; use of data-processing equipment, 154; use of ratios, 153

Orientation and special guidance courses: 209

Parent groups, counseling with: 208

Philosophical trends in guidance, counseling, and personnel services: 146

Preparation and professionalization of other specialists: 175

Preparation of counselors: 172

Professionalization of counselors in educational settings: 174

Program evaluation: control group studies, 216; criterion problem in, 218; follow-up studies, 215; general discussions, 214; methodological problems in, 218; problems of experimental design in, 220

Roles and functions of pupil personnel workers in relation to: health services, 156; nomenclature of guidance services, 157; psychological services, 157; role and function of the counselor, 158; self-perceptions of counselors, 157

Selection of personnel specialists: 171
Selection, preparation, and professionalization of specialists: 171
Society and culture in guidance, counseling, and personnel services in relation to: college students, 143; culturally deprived, 145; school-age youth, 144; values, 144

Test taking and its preparation in relation to: instructions, 188; perception, 189; practice, 188; screening, 189

Vocational development process: career development, 197; occupational classification, 200; occupational information, sources of, 201; occupational psychology, 199; women's career development, 200
Vocational guidance in regular course work, 210

REVIEW OF EDUCATIONAL RESEARCH

Official Publication of the American Educational Research Association
Contents are listed in the Education Index.

Copyright 1963
By National Education Association of the United States, Washington, D.C.

Volume XXXIII, No. 3 **June 1963**

Curriculum Planning and Development

Reviews the literature for the three-year period since the issuance of Vol. XXX, No. 3, June 1960

TABLE OF CONTENTS

Chapter *Page*

Foreword 230

I. Chart or Charter: Recent Developments in Educational Discourse 233
DAVID L. ELLIOTT, *Teachers College, Columbia University, New York, New York*
ARTHUR W. FOSHAY, *Teachers College, Columbia University, New York, New York*

II. Organized Knowledge Influencing Curriculum Decisions...... 245
O. L. DAVIS, JR., *Kent State University, Kent, Ohio*

III. Societal Forces Influencing Curriculum Decisions............ 254
GORDON F. VARS, *Cornell University, Ithaca, New York*
WILLIAM T. LOWE, *Cornell University, Ithaca, New York*

IV. Psychological Forces Influencing Curriculum Decisions....... 268
NORMAN D. BOWERS, *Northwestern University, Evanston, Illinois*

V. Curriculum Components and Organization.................. 278
RICHARD A. GIBBONEY, *Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania*
JOHN H. SANDBERG, *Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania*
DONALD M. CARROLL, *Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania*
STANLEY N. MILLER, *Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania*
NORMAN A. MILLER, *Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania*

Chapter Page

VI. Teaching 293

ROLLAND CALLAWAY, *University of Wisconsin—Milwaukee, Milwaukee, Wisconsin*
GERALD T. GLEASON, *University of Wisconsin—Milwaukee, Milwaukee, Wisconsin*

VII. Curriculum and Communications Media.................... 301

CURTIS PAUL RAMSEY, *George Peabody College for Teachers, Nashville, Tennessee*

VIII. Curriculum Development: Dynamics of Change............. 304

PAUL R. KLOHR, *Ohio State University, Columbus, Ohio*
JACK R. FRYMIER, *Ohio State University, Columbus, Ohio*

IX. Curriculum Research: Problems, Techniques, and Prospects... 322

JAMES B. MACDONALD, *University of Wisconsin—Milwaukee, Milwaukee, Wisconsin*
JAMES D. RATHS, *University of Wisconsin—Milwaukee, Milwaukee, Wisconsin*

X. Curriculum Planning and Development: A Review of Reviews.. 330

WILLIAM M. ALEXANDER, *George Peabody College for Teachers, Nashville, Tennessee*

Erratum

Volume XXXIII, No. 1, February 1963, page 127, note b to Table 1: *For* L. S. Blackman *read* L. S. Blackman and M. P. Smith, Co-Principal Investigators; R. Capobianco, Assistant Director

This issue of the REVIEW was prepared by the
Committee on Curriculum Development and Planning

O. L. DAVIS, JR., *Chairman*, Kent State University, Kent, Ohio
WILLIAM M. ALEXANDER, George Peabody College for Teachers, Nashville, Tennessee
NORMAN D. BOWERS, Northwestern University, Evanston, Illinois
ARTHUR W. FOSHAY, Teachers College, Columbia University, New York, New York
RICHARD A. GIBBONEY, Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania
GERALD T. GLEASON, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin
PAUL R. KLOHR, Ohio State University, Columbus, Ohio
JAMES B. MACDONALD, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin
CURTIS PAUL RAMSEY, George Peabody College for Teachers, Nashville, Tennessee
GORDON F. VARS, Cornell University, Ithaca, New York

with the assistance of

ROLLAND CALLAWAY, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin
DONALD M. CARROLL, Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania
DAVID L. ELLIOTT, Teachers College, Columbia University, New York, New York
JACK R. FRYMIER, Ohio State University, Columbus, Ohio
WILLIAM T. LOWE, Cornell University, Ithaca, New York
NORMAN A. MILLER, Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania
STANLEY N. MILLER, Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania
JAMES D. RATHS, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin
JOHN H. SANDBERG, Pennsylvania State Department of Public Instruction, Harrisburg, Pennsylvania

FOREWORD

The first issue of the REVIEW's first volume was devoted to the curriculum; the present issue is the eleventh to be so concerned. During the 32 years covered by this series, the issues not only have mirrored the advocated and actual emphases of the American school curriculum but also have provided valuable assistance in advancing both the substance and methodology of research on the curriculum. The committee which prepared this issue sincerely hopes that it may contribute to the distinguished series of which it is a part and may serve to further the cause of curriculum research and theory construction.

While this issue continues to give attention to several persisting curriculum topics, many noticeable changes have been made in its organization and in its treatment of research studies selected for review. Three new chapters have been included, but, perhaps more important, the attention of others has been refocused. Such alterations in organization were made so that the issue might deal comprehensively with the curriculum field and specifically with important elements of curriculum development and planning. Conscious efforts were made to minimize overlap between chapters in this issue and those chapters in previous issues of the REVIEW. Where duplication of topics occurs, the concern of chapters in this issue is specific to the curriculum and not to the topics in general. Priority for review was given to evaluative and comparative research studies whenever possible.

Selection of studies to review in this chapter was, in many cases, not an easy task. The most important factor explaining this difficulty was the decision to insist that studies be reviewed, not merely listed or annotated. To implement this decision, no more than 60 sources were reviewed in any one chapter. Most of the research writings selected for review were considered more relevant and, hopefully, more significant than others related to the topic of a particular chapter.

The committee which developed this issue of the REVIEW echoes its predecessors in lamenting the paucity of well-conceived, well-designed, and rigorously analyzed studies in the field of curriculum. Nevertheless, it dissents from previous estimates about the probable continuation of this situation. The past three years appear to this committee to hold the prospect of important, if not dramatic, changes in the curriculum research scene. Among the reasons for optimism are a growing concern for research and study of truly significant problems, a developing sophistication in research design and analysis, and an awareness of the relevance to curriculum research of models and of techniques originated in other theoretical and applied fields of inquiry. Added to these conditions is the increasing commitment of money, time, and resources, along with encouragement to engage in curriculum research.

The chairman wishes to thank the members of the committee and the contributors for their competent cooperation in this venture. Beset by many problems, individuals conscientiously worked on assignments, at times against formidable obstacles and at personal sacrifice. Special appreciation is accorded the REVIEW editor who contributed so much to whatever success this issue may have.

O. L. Davis, Jr., *Chairman,*
Committee on
Curriculum Development and Planning

CHAPTER I

Chart or Charter: Recent Developments in Educational Discourse

DAVID L. ELLIOTT and ARTHUR W. FOSHAY

THE DISCOURSE about educational matters—and especially about curriculum—continued at an accelerated pace during the past three years. One reason for the vast quantity of publications with emotional overtures is that education generally, and curriculum particularly, remain uncharted fields in which the function of scientific inquiry has been questioned at various times and in which the relationship between what ought to be obtained and what can be done is essentially unexamined. The statements with the greatest public impact are charters, creeds, and manifestoes. Meanwhile, the brotherhood represented by readers of the REVIEW continues to act as if a rational examination of the hopes and possibilities expressed in the professional literature could yield hypotheses which, when tested, would yield a realistic basis for modifications in or for development of theories about the educational process.

This chapter is based on the assumption that improved theories of education are desirable. It also assumes that many of the abusive criticisms and emotionally charged discussions arise because no coherent theory exists. If this latter assumption is defensible, then the attempts at theory making as well as the renewed criticism of the present state of educational research can be seen as having considerable significance.

Criticisms of Educational Research

Educational research has been criticized in many ways. Many have criticized it for the kinds of problems tackled; several, for the procedures used; and some, for the way in which the results are, or are not, employed. Research directed at educational practice is suspect for many, not only because it involves applying scientific methods to human beings engaged in value-directed activities but also because it lacks rigor. Perhaps the most common complaint is that research knowledge in education does not cumulate in the manner found in such fields as medicine or engineering.

A brief survey of the literature is illustrative. Coladarci (1960), Jensen (1962), and Joncich (1962) pointed to a lack of sophistication on the part both of researchers and of those who employ their findings. Joined by Belth (1962), Goodlad (1962), and Millar (1960), both Carroll (1961) and Coladarci (1960) urged that emphases be placed upon pertinent but central research problems of a varied nature. Beauchamp (1961), Coladarci (1960), Conant (1961), Maccia (1962b), and Ward (1961) saw a need for a firm theoretical grounding for research.

Generally representative of the trend of criticism are articles by Joncich (1962) and Tyler (1962). Joncich (1962) traced the rise of a scientific movement in educational research and the subsequent abandonment of the idea of a science of education during the first quarter of this century. She characterized research efforts since the 1930's as being overly analytic and fragmentary and exhibiting too little of the purpose, vigor, and "daring hypotheses" of the earlier period. Joncich concluded that education could and must assume the properties of a rational enterprise that is anchored to a clearly comprehensive and objectively derived theoretical framework.

Tyler (1962) indicated the severe limitations of local evaluations as a source of useful generalizations and stressed the need for basic research, that is, research which, through developing and testing a theory, could serve to explain educational phenomena and to guide the development of more effective educational practices and materials. Conant (1961) made a distinction between trial and error experimentation which builds up practical knowledge (e.g., recipes for particular situations) and experimentation based upon a theoretical component which gives rise to generalizations (e.g., laws applicable to a wide variety of situations). Upon such generalizations, explanations could be based and predictions could be deduced.

In this same spirit, Hearn (1958) spoke of social work as a "science-based art," neither wholly dependent on the personal artistry of the social worker nor completely included in any of the social sciences. Characterized in this way, social work—or, by analogy, education—is, however, at least partially susceptible to empirical study and theoretical formulation. It is not unthinkable that the artistic aspects of teaching or leadership might be made clearer and more familiar to the public by employing the analytic and synthetic tools of literature, the fine arts, and aesthetics. For example, Bantock (1961) suggested that novelists might help the social scientists to identify key components in and the feeling underlying typical social situations.

Nature of Research Theory

Three Stages of Inquiry

It seems apparent, however, that confusion exists as to what theory is and how it applies to education. Northrop (1959) suggested that three stages characterized fruitful inquiry in any field: (a) first, a phase embodying an analysis of the problem in which the basic elements are discerned; (b) second, a natural history effort embracing observation, description, and classification; and (c) third, an attempt at deductive formulation of a theory—a process in which inferences are added to facts in order to obtain concepts and testable hypotheses. These stages are not necessarily apparent in the work of any one research worker, since he may be standing on the shoulders of others in his field. Moreover, the order in which the three stages are pursued may not be even crucial, provided that each one is followed through thoroughly.

To the extent that the concepts used in research contain inferences, they depend for their meaning—and thus, ultimately, for their usefulness in public inquiry—upon the theoretical point of view of those who do the inferring. If that point of view is a well-developed, explanatory, predictive theory, then the concepts within its domain are defined both denotatively and connotatively—both for themselves and for their inner relationships with other concepts in the same domain. Such concepts are known by the theorists who postulate them. It follows, then, that to borrow a concept from one domain to use in another, without bringing its theoretical roots along, is to cut it off from much of its meaning and to curb its usefulness. The terms *student, story,* or *teacher* may mean the same thing to many people in many contexts; the terms *learner, print media,* and *programed instruction* depend for their meaning upon particular theoretical contexts.

Analysis of Problems

Many suggestions for improving educational research seem to fit into Northrop's first stage of analysis of the problem. Carroll (1961) asked both for a central focus on the school learning process with a time dimension added and for close attention to actual behavior and real-life situations. He was joined by Millar (1960), who emphasized the subject matter disciplines; by Jensen (1962), who stressed a need for careful study of the hereditary, developmental, and environmental factors conducive to learning; and by Ausubel, who suggested in the December 1961 REVIEW that another careful look at verbal reception learning be taken in the course of redoubled efforts to try to locate what it is that makes any learning meaningful. Also related were articles by Goodlad (1962) on subject matter organizing centers, which might provide for a range of student responses, and by Bantock (1961), who suggested a focus on different kinds of teaching in varying contexts.

Educational philosophers, like Northrop, play an important part in this analytic stage. Conant (1961), Hardie (1962), and Maccia (1962c) join Northrop in urging the necessity of making distinctions between normative problems involving values and factual problems.

Hardie (1962), Nagel (1960), and Scheffler (1960) analyzed selected educational terminology relative to contextual and logical considerations. Specifically, both Nagel (1960) and Hardie (1962) called into question some assumptions underlying measurement.

Page (1962) found a "verbal magic" rather than careful attention to the realities of school situations present in many of the attempts to relate behavioral science findings to educational practice. Conclusions drawn from experiments with rats, for example, are often applied directly to work with school students without the necessary mediating inquiry. To remedy this, Page suggested three courses of action: (a) analyzing the physical realities of experimental situations, (b) harnessing excellent design and measurement skills to education-centered problems, and (c) expressing these problems in language shaped for education which is "intersubjectively testable and scientifically respectable."

Finally, Smith (1961) employed philosophical and analytical techniques to the concept of teaching and to its relationship to learning. He concluded that teaching should be examined as a discrete entity which, in being independent both of concepts of process and of specific conditions, would connote its own forms, constituent elements, and regularities. Upon this conclusion, Smith and others have based large-scale research projects.

Natural History Stage

The movement from an immediately perceived to a described fact in Northrop's natural history stage is illustrated in a project reported by Hemphill, Griffiths, and Frederiksen (1962). In a phase of their work entailing the building of concepts for use in research on educational administration, they proceeded in three steps. First, they drew upon material that described the job of an elementary principal. The summary of this material resulted in a taxonomic grid with 12 cells containing four areas of administrative activity taking place on three levels. Second, incidents were created that might occur in any one of the boxes in the grid, that is, examples of activity in each area and at each level of administrative activity. From these examples, realistic facsimiles of problems with which principals might actually have to deal were devised. These items were presented to principals in simulated situations which, among other procedures, employed "in basket" tests and kinescopes of school situations. Finally, the data collected were scored in several ways and factor-analyzed. The factor analysis yielded eight basic factors or dimensions of administrative performance. It was pointed out that the categories had practically no relationship in language to the categories of the original taxonomy. This was a systematic analysis of a process or function that went from common sense terminology to concepts intended to be more useful in describing the phenomena of educational administration.

Several taxonomic schemes appear in the literature. These are variously called conceptual frameworks, conceptual systems, and models. (The term *model* is also used in certain special ways, which will be discussed later.) Since a taxonomy is not a theory but an arrangement of the basic elements, it represents essentially Northrop's natural history stage, which may lead to a theory. It should be emphasized that the chief shortcoming of the various curriculum frameworks arises from the difficulty that they are not related to inferences, hypotheses, or action. As curriculum guides, such frameworks cannot function. They are pretheoretical and, therefore, two great steps away from action.

Downey (1960) presented a "model for the improvement of secondary education," which he had designed (a) to systemize existing knowledge about that field into three dimensions of the educational enterprise—substantive, procedural, and environmental; (b) to point up the need for further research and knowledge; and (c) to provide a plan for the improvement of practice. Millar (1960) suggested that a conceptual system, or framework, serves an educational researcher by aiding in the identification of important problems. An adequate conceptual system should (a) aid in the finding of necessary relationships between various problems in order that the organization of the framework may be intrinsic to the problems involved; (b) help suggest data sources; (c) serve not only to identify points relative to which curriculum decisions have to be made but also to designate resources by which such decisions may be made and implemented; and (d) supply criteria by which existing research results may be evaluated. In addition to all this, an adequate framework is self-correcting and, thus, modifiable in the light of new knowledge and changing conditions.

Beauchamp (1961) saw what he called a "curriculum theory model" as being one step beyond taxonomy on the way to construction of curriculum theory and subtheories. He stated that such a model contains assumptions and postulates, both normative and descriptive, that are basic to curriculum and curriculum development, definitions of technical terms, statements about ways to involve people in curriculum planning, and prescriptions for certain curriculum design features. The components of such a model are chosen on the basis of a classification system or taxonomy of the elements contained in curriculum development work. The specific factors included in the model determine its nature and the nature of the theory produced. The model is used as a guide to curriculum choices. When such choices are expressed, Beauchamp believes that a curriculum theory has been formulated.

Theory for Education

Two Usages of the Term

The term *theory* has been employed in the literature in two main senses, and, in keeping with the emphasis of this chapter, a definite distinction must be made between the two. The first usage which falls within the domain of the philosophy of education as described by Clayton in the February 1961 REVIEW or Bayles (1962) is exemplified by an entire issue of *Educational Theory* (1961)—an issue dealing with certain philosophical statements or syntheses by Alfred North Whitehead and with their application to educational practice. Philosophies of education may include statements about the nature of man, society, and knowledge as well as normative statements concerning the proper aims and goals of schooling. Such philosophies are not necessarily consistent with the descriptive findings and theories of the behavioral sciences. Moreover, deductions directly applicable to actual school practice frequently cannot be made from statements of aims.

The second usage is illustrated by a group of articles in *Educational Theory* by Ausubel (1961) on verbal learning, by Page (1962) on experimental models, by Belth (1962) on the prospects for a discipline of education, and by Clements (1962) on the use of the term *theory* in education.

Theory in Relation to Curriculum Development

During the past three years, several writers have stressed the importance of theory as the basis for carrying on educational research. Beauchamp (1961) presented perhaps the most ambitious treatment of theory relevant to curriculum development. In his book, he reviewed a number of statements about theory in the behavioral sciences and then proposed his own definition of curriculum theory. According to Beauchamp, theoretical work in curriculum is concerned with definition of technical terms, classification of existing knowledge, formulation of inferences and predictions, design of curriculum theory models to guide theory construction, and development of subtheories, since the immensity of the curriculum field prevents its being investigated as a whole. This subtheory category includes the interpreting of theories from other disciplines such as psychology for curriculum, raising questions of curriculum design, seeking improved answers, and devising theory for curriculum design and other aspects of curriculum. Beauchamp concluded by offering a curriculum theory model of his own, both as an example of what he had described and as a start for further work by others. In the terms used here, Beauchamp's model, which is still a framework, also includes normative assumptions.

Theories in Educational Research

Tyler (1960) characterized behavioral science theories as conceptualizations revealing relationships, structures, and processes which are embedded in complex situations. The most useful of these theories not only guides an observer in grasping more readily the events in a classroom, for example, but also permits explanation of such events. In a later article, Tyler (1962) noted that theories could function in two main ways in educational research: (a) as bases for explanation of educational phenomena, and (b) as sensible guides for the invention of procedures and materials for educational practice. The theory aids one in deciding what to look for and how to look for it.

Maccia (1962b) described a theory as a representative model of the phenomena in a particular domain and as a set of statements at the level of conjectures that characterize the objects in the domain. Maccia and Maccia (1962) indicated that an adequate theory and characterization of aspects of the educative process are needed.

Other writers seem to be in general agreement with the preceding statements on the function of theory for educational research. However, there seems to be a great deal of disagreement and confusion concerning (a) what aspects of the educational enterprise are accessible to theoretical formulations of the behavioral science type and (b) how one should go about obtaining or constructing a theory. The impression that one gains from Beauchamp's suggestions is very different from that which one obtains from the examples given by Tyler (1962), Maccia (1962a), and Griffiths (1959, 1963).

Theory Defined

At this point in the history and development of educational research, it seems well to start with a rigorous and narrow definition of descriptive theory. In his helpful discussion of administrative theory, Griffiths (1959) accepted a restricted definition by viewing a theory as furnishing a number of assumptions from which a set of unifying empirical laws may be derived by logical procedures. Similarly, Hearn (1958) proposed that a theory consists of a number of verified hypotheses that are internally consistent.

Internal consistency is what makes the separate propositions a set. The definition of a theory can be tightened further by suggesting that (a) no one statement in the set should be derivable from any other statement or combination of statements in the set, and (b) the total set of laws cannot be derived without the inclusion of every member of the set. A theory is tested through the hypotheses or laws that are deduced from it. Thus, one criterion of adequate descriptive theory is that its laws are logically, or in principle, verifiable even though the necessary instruments or techniques may not have been invented yet.

From Tyler's (1960) formulation, one could say that theories for education would help educational researchers to select or abstract from complex situations certain variables that are systematically related to each other. Each set of related variables, or concepts, constitutes a system; and the theory provides, in Griffiths' (1959) formulation, not only a description of data sets in terms of symbols but also the means for performance of operations upon the symbols in line with rules underlying the system of which the symbols are a part. The rules of a theoretical system are its laws and the assumptions that explain the laws. The concepts postulated and defined by the theory are the variables contained in the system. These variables behave according to the rules or laws or, if not verified, according to the hypotheses proposed. Laws and hypotheses are "if-then" statements about the regularities of the interrelationships of the variables. To the extent that the concepts employed in a research design are part of the same theoretical system, they can be fruitfully manipulated in varying relationships to each other, and the conclusions or results will be generalizable to other situations in which the same or analogous systems of concepts or variables are to be found.

Through covering only a limited domain, a theory helps the researcher to delimit the scope of his research effort. The limited scope in turn is a subdivision of the educational enterprise, such as teaching or part of teaching, leadership, learning, or group behavior. It seems best, therefore, first to seek theories for education or curriculum—distinct theories of distinct aspects of the educational enterprise. Later, a comprehensive descriptive theory of teaching or a comprehensive descriptive-normative curriculum theory or philosophy may follow.

Theory Making in Education

In the literature of the past several years, two main approaches to the building of behavioral science theory for education at a descriptive level were proposed. The first is inductive, proceeding from observable data and working up as far as possible from hypotheses to laws and then to explanatory assumptions. The second is deductive, or, in Maccia's (1962a, b) terminology, retroductive, in that it involves borrowing theoretical formulations from other analogous areas of inquiry, substituting concepts, deducing hypotheses and laws, and testing the results against observable data.

The inductive approach to theory building, which is found in the previously mentioned stages of inquiry suggested by Northrop (1959), can be related to a set of rules for theory building worked out by Homans (1950) as well as to the content of aforementioned writings of Beauchamp (1961), Hearn (1958), and Hemphill, Griffiths, and Frederiksen (1962).

A similar approach was used in the study of intellectual functioning by Jean Piaget, as described by Hunt (1961). Piaget observed the behavior of children at different ages when they were presented with varying kinds of intellectual tasks. He made theoretical inferences about the presence and development of certain intellectual structures which might account for what was observed.

The deductive approach was most completely stated in the writings of Griffiths (1959, 1963), Maccia (1962a, c), Maccia and Maccia (1962), and Tyler (1960).

Models

Griffiths (1963) proposed an approach to theory making through the use of models. As did Tyler (1960), Griffiths suggested the use of the term *model* to denote a theoretical formulation from an area about which a great deal is known. The available information is used to suggest laws, or hypotheses, for an area about which little is known. Both pointed to the futility of using formulations from areas about which little is known to study other areas about which little is known. Both Griffiths (1963) and Maccia (1962b) illustrated this operation of theory development through an example taken from Brodbeck (1959) concerning the problem of whether the laws of epidemiology could be used as a model for a theory of rumor transmission. In Brodbeck's example, the descriptive concepts of the laws of epidemiology were first replaced by letter variables to reveal the form of the laws. The concepts referring to diseases were then put into a one-to-one correspondence with those referring to rumors, and the letter variables of the epidemiology laws were replaced by the descriptive terms referring to rumors. This resulted in a set of hypotheses about rumors, which may or may not be confirmed.

Griffiths went on to identify some of the assumptions that underlie the use of models. For the study of educational administration, Griffiths suggested that an open systems model could be applied to the dimensions of administrative performance identified through factor analysis, as in the study reported by Hemphill, Griffiths, and Frederiksen (1962). Two assumptions about educational administration that would be involved in making this application are that (a) administrative performance is a generalized type of behavior found in all human organizations, and (b) administrative performance involves directing and controlling behavior in a social enterprise.

In applying a model to a domain under study, one looks for isomorphism, that is, both for a one-to-one relationship between the elements of the model and the elements of the domain and for the same kind of relationship to one another among the elements of the model and the elements of the domain. Extended to theory, this statement can be interpreted to mean that, if correspondence in the form of two sets of laws in two theories exists, one theory is a model for the other. The propositions derived from the model theory must be demonstrated as applicable to the aspect of education under study.

In using the term *retroduction* for the process of letting one theory furnish a basis or model from which another theory could be set forth, Maccia (1962b, c) seemed to differ with Griffiths on the extent to which isomorphism or analogy is desirable. Following an analysis and discussion of induction and deduction, she said that, since neither induction nor deduction covers the process involved in devising theory, together they do not completely describe scientific inquiry. Maccia (1962a) concluded that three principal stages of scientific inquiry exist: (a) retroduction, through which one may set up characterizations; (b) deduction, through which one can explicate characterizations; and (c) induction, through which one may evaluate the characterizations.

Incomplete Analogy

An important aim in education is to devise new theory, that is, to formulate characterizations not already possessed about the domain or aspect of education under study. If Maccia's main point has been interpreted correctly, she differs with Griffiths in saying that in order for new theory or fresh characterizations to be fruitful, the components of the theory used as a devising model cannot be completely analogous to components of the areas under study. If they are strictly analogous, then the devising model may be already an adequate theory for the aspect of education under study. If the model is not properly called a theory, the terms *framework* and *conceptual scheme* may be preferable.

Maccia (1962a, b) made her point about the nature of devising models by classifying models into several distinct categories. The two major categories are the representational model, or a model *of* a particular set of phenomena or domain, and the nonrepresentational model, or a model *for* a particular set of phenomena. A theory that operates for a domain is a representational model, as in Tyler's description of theory as a simplified picture that characterizes a particular set of phenomena and their interrelationships. Although frameworks are also representational models, they differ from theory in other respects.

In a survey of recent literature dealing with the use of models for the study of various aspects of education, Maccia (1962c) classified models as either representational, thus promising for explication and evaluating theory, or nonrepresentational, hence promising for devising theory. She identified those of the 37 models included in the survey that seemed to fulfill the criteria for theory-devising models. Among those studies identified as suggesting models of the nonrepresentational, characterization type were the following: (a) Ausubel's exposition of a programmatic model for systematically ordering relationships, offered as a plan in the December 1961 issue of the REVIEW; (b) Carroll's (1959) formulation of a model for language behavior based on computer logic; (c) Griffiths' (1963) presentation of a characterization theory which was used to devise another characterization theory; (d) Kjeldergaard's description, in the April 1961 REVIEW, of a theory of mediated generalization, in which a computer program was used to devise a theory of the acquisition of language; and (e) Sweitzer's (1963) proposal that two theories can be used not only to explicate and perhaps to evaluate each other but also to devise further propositions not contained in either theory.

Finally, one can conclude that there is a third major source of the descriptive type of theory to serve and direct educational research. The first two sources involve the analysis and description of peculiarly educational phenomena or patterns of phenomena. The third entails the employment of theories of human behavior from the several behavioral sciences in the manner illustrated by Tyler (1960). The aspects of human behavior encompassed by theories in the behavioral sciences are common to all human individuals or groups of individuals. Although the educational researchers and practitioners may make particular applications of such theories to the educational domain, there is assumed to be no exclusively educational theory underlying the behavior in question. By way of example, this idea seems inherent in Griffiths' (1963) discussion of the application of systems theory to educational administration.

Whereas a systems theory of administration could be employed to study the administrative function in any organization and a social psychology theory of group dynamics could be applied in the study of many different kinds of groups, there are areas in the educational enterprise, e.g., aspects of teaching or being a student, for which no theory from any other discipline exists per se. Thus, it seems important to bring to bear Northrop's three stages of inquiry on these two areas for which no adequate theory has been found or devised. One needs also to follow up Maccia's conception of retroduction through nonrepresentational models—an important development in theory construction which has emerged from the Center for the Development of Theory in Education at Ohio State University. The use of a point of view from within a theory in one field as a path leading to fresh characterizations or theoretical formulations in another field has apparently been the key to progress in much scientific inquiry in the past; the concept *nonrepresentational* may be a useful criterion of selection in one's search for theories to use in devising models for educational theory.

Bibliography

AUSUBEL, DAVID P. "In Defense of Verbal Learning." *Educational Theory* 11: 15-25; January 1961.

BANTOCK, G. H. "Educational Research: A Criticism." *Harvard Educational Review* 31: 264-80; Summer 1961.

BAYLES, ERNEST G. "Educational Imperatives." *Educational Theory* 12: 129-39; July 1962.

BEAUCHAMP, GEORGE A. *Curriculum Theory*. Wilmette, Ill.: Kagg Press, 1961. 149 pp.

BELTH, MARC. "Prospects for a Discipline of Education." *Educational Theory* 12: 193-204; October 1962.

BRODBECK, MAY. "Models, Meaning, and Theories." *Symposium on Sociological Theory*. (Edited by Llewellyn Gross.) Evanston, Ill.: Row, Peterson and Co., 1959. Chapter 12, pp. 373-403.

CARROLL, JOHN B. "An Operational Model for Language Behavior." *Anthropological Linguistics* 1: 37-54; January 1959.

CARROLL, JOHN B. "Neglected Areas in Educational Research." *Phi Delta Kappan* 42: 339-43; May 1961.

CLEMENTS, MILLARD. "Theory and Education." *Educational Theory* 12: 124-28; April 1962.

COLADARCI, ARTHUR. "Towards More Rigorous Educational Research." *Harvard Educational Review* 30: 3-11; Winter 1960.

CONANT, JAMES B. *Trial and Error in the Improvement of Education*. Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1961. 24 pp.

DOWNEY, LAWRENCE W. "Secondary Education: A Model for Improvement." *School Review* 68: 251-65; Autumn 1960.

EDUCATIONAL THEORY. "Centennial Issue: Alfred North Whitehead, 1861-1947." *Educational Theory* 11: 193-280; October 1961.

Goodlad, John I. "The Organizing Center in Curriculum Theory and Practice." *Theory into Practice* 1: 215-21; October 1962.

Griffiths, Daniel E. *Administrative Theory*. New York: Bureau of Publications, Teachers College, Columbia University, 1959. 123 pp.

Griffiths, Daniel E. "Some Assumptions Underlying the Use of Models in Research." *Educational Research: New Perspectives.* (Edited by Jack A. Culbertson and Stephen P. Hencley.) Danville, Ill.: Interstate Printers and Publishers, 1963. Chapter 9, pp. 121-39.

Hardie, Charles D. *Truth and Fallacy in Educational Theory*. American edition. New York: Bureau of Publications, Teachers College, Columbia University, 1962. 156 pp.

Hearn, Gordon. *Theory Building in Social Work.* Toronto: University of Toronto Press, 1958. 83 pp.

Hemphill, John K.; Griffiths, Daniel E.; and Frederiksen, Norman. *Administrative Performance and Personality: A Study of the Principal in a Simulated Elementary School.* New York: Bureau of Publications, Teachers College, Columbia University, 1962. 432 pp.

Homans, George C. *The Human Group.* New York: Harcourt, Brace and Co., 1950. 484 pp.

Hunt, Joseph McVicker. *Intelligence and Experience.* New York: Ronald Press, 1961. 416 pp.

Jensen, Arthur R. "The Improvement of Educational Research." *Teachers College Record* 64: 20-27; October 1962.

Joncich, Geraldine. "Wither Thou, Educational Scientist?" *Teachers College Record* 64: 1-12; October 1962.

Maccia, Elizabeth Steiner. *The Conceptions of Model in Educational Theorizing.* Occasional Paper 62-114. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 53 pp. (Mimeo.) (a)

Maccia, Elizabeth Steiner. *Models and the Meaning of "Retroduction."* Publication 62-110. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 16 pp. (Mimeo.) (b)

Maccia, Elizabeth Steiner. *Ways of Inquiring.* Publication 62-107. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 13 pp. (Mimeo.) (c)

Maccia, Elizabeth Steiner, and Maccia, George S. *The Way of Educational Theorizing Through Models.* Publication 62-111. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 16 pp. (Mimeo.)

Millar, Michael. "Some Roles of Subject Matter in Curriculum Inquiry." *School Review* 68: 329-45; Autumn 1960.

Nagel, Ernest. "Philosophy in Educational Research." *First Annual Symposium on Educational Research.* (Edited by Frank W. Banghart.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1960. Chapter 5, pp. 71-84.

Northrop, F. S. C. *The Logic of the Sciences and the Humanities.* New York: Meridian Books, 1959. 402 pp.

Page, Ellis B. "Behavioral Theory, Verbal Magic, and Education." *Educational Theory* 12: 73-78; April 1962.

Scheffler, Israel. *The Language of Education.* Springfield, Ill.: Charles C. Thomas, 1960. 113 pp.

Smith, B. Othanel. "A Concept of Teaching." *Language and Concepts in Education.* (Edited by B. Othanel Smith and Robert H. Ennis.) Chicago: Rand McNally & Co., 1961. Chapter 6, pp. 86-101.

Sweitzer, Robert E. "An Assessment of Selected Theories of Administration." *Educational Research: New Perspectives.* (Edited by Jack A. Culbertson and Stephen P. Hencley.) Danville, Ill.: Interstate Printers and Publishers, 1963. Chapter 13, pp. 199-231.

Tyler, Ralph W. "The Contribution of the Behavioral Sciences to Educational Research." *First Annual Symposium on Educational Research.* (Edited by Frank W. Banghart.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1960. Chapter 4, pp. 55-70.

Tyler, Ralph W. "Specific Contributions of Research to Education." *Theory into Practice* 1: 75-80; April 1962.

Ward, Virgil S. "The Function of Theory in Programs for the Gifted." *Teachers College Record* 62: 532-39; April 1961.

CHAPTER II

Organized Knowledge Influencing Curriculum Decisions

O. L. DAVIS, JR.

SERIOUS examination of the nature of knowledge and of its proper relationship to the school curriculum has been evident during the past three years. From the resulting analyses and ancillary activities has emerged the rapidly developing recognition that organized knowledge is a major source of the curriculum. Thus, the nature of the intellectual disciplines and of the practical technologies demands as careful and systematic attention in curriculum development and planning as the nature of society and the learner received in the past. This response undoubtedly was influenced by multiple pressures, including (a) the recognition that the disciplines are instructive in addition to preservative in nature, (b) the renewal of interest and active participation in curriculum planning by academic scholars, (c) the expression of general concern throughout American society for increased excellence in schooling, and (d) the determination to deal effectively with the modern explosion of knowledge.

Developments of the recent past and prospects for the future, rather than precedent or a fund of research evidence, justify the inclusion in this issue of the REVIEW of a chapter dealing with organized knowledge and the curriculum. In this chapter, some of the recent analytical statements and descriptive reports treating the nature of knowledge and the curriculum are reviewed. A major purpose is to highlight those statements and reports felt to be particularly productive in terms of future development and research. Not attempted here is a comprehensive survey of attempts to analyze most disciplines for their curriculum implications, a historical treatment of the role of organized knowledge in past curriculum-making formulations and efforts, or an analysis of the conceptions of knowledge revealed in recently inaugurated curriculum development projects.

The Disciplines: A Curriculum Source Recognized

In his stimulating and influential book, *The Process of Education*, Bruner (1960) emphasized the importance of the structure of knowledge in curriculum planning and teaching. He pointed out that to know the structure of a subject is to know its patterns and concepts. To deal with the fundamental structure of a field, he held, will make the subject understandable, facilitate transfer and retention of that which is learned, and reduce the lag between knowledge dealt with in school and that produced by scholars. Bruner gave special attention to the importance of curriculum provisions for both intuitive and analytic thinking. His emphasis on the desirability of fostering intuition was consistent with his provocative position on the structure of knowledge. Bruner's (1962) more recent book, *On Knowing*, presented elaborations of his basic positions regarding structure of knowledge and thinking. The "lucky guessing" which he advocated should not be confused with intellectual anarchy: Discovery, which is based on one's having prior knowledge with which to think, consists of reordering data in ways that permit one to go beyond the data to new understanding. Bruner did not see his papers as establishing the basic criteria for curriculum selection and organization. He presented, in a popular and engaging manner, informed and competent judgments on topics necessarily important in curriculum planning. Since he did not prescribe comprehensive solutions, curriculum workers and researchers must specify and evaluate applications derived from his general propositions.

Foshay's (1961) address to the conference of the Association for Supervision and Curriculum Development claimed for organized knowledge—the disciplines and technologies—overdue recognition as a basic curriculum source. Each discipline, he asserted, is characterized by three elements: its domain, its way of knowing or method of inquiry, and its tradition. To approach knowledge indirectly, as he believed curriculum making has done in the past, is to take knowledge for granted. His modest proposal was a direct confrontation in the schools of the essence of the disciplines. In addition to his illuminating examples of direct approaches to the disciplines, Foshay discussed educators' roles in translating the disciplines into subject matter appropriate for learning by school children. An indispensable ingredient in this translation task was seen as the educators' insisting that scholars effectively communicate their conceptions of their disciplines. For such communication to take place, Foshay urged curriculum workers to ask the right questions, to listen attentively, and then to develop programs using the insights thus gained in addition to understandings of society and learners. Foshay also insisted that direct attention to the disciplines must not obscure or replace concern for important, complex problems to which the contributions of any single discipline would be inadequate.

Phenix (1962) stated that only knowledge drawn from the disciplines is appropriate to the curriculum. His cogent argument posited that knowledge in a discipline is instructive and, while knowledge outside the disciplines exists, it is noninstructive. Phenix indicated that disciplined knowledge is characterized by analytic simplification, synthetic coordination, and dynamism. Also provocative was his contention that the disparity between the logic of the disciplines and the psychologic of instruction is imagined and destructive to both scholarship and pedagogy. Phenix also called for acceptance of a realistic rather than a nominalistic view of knowledge for curriculum planning.

The Disciplines: Analyses and Implications

Two types of positive responses have followed the call for greater attention to the disciplines in curriculum development. Several important attempts have been made to analyze particular disciplines and their structures. The significance of these efforts lies not in that they were made, for some scholars throughout time have studied and discussed the nature of the disciplines. Rather, their importance is derived from the fact that they were made in the hope that clearer, more accurate, and more coherent curriculums might be constructed for elementary and secondary schools. Another response has been the formulation of proposals to take into account the various disciplines within a general curriculum design.

A major contribution to the necessary understanding of the disciplines was a seminar convened in 1961 by the National Education Association, Project on Instruction. At this meeting, scholars discussed with educators fundamental concepts and inquiry modes from their disciplines. The seminar report (National Education Association, Project on Instruction, 1962) summarized the basic ideas treated and suggestions advanced regarding the relationship of the disciplines and the school curriculum. One interesting proposal was a curriculum with "nuclear" and "cortical" components. Without some such curriculum arrangement, some participants foretold either an internecine power struggle between proponents of individual disciplines or a possible homogenizing synthesis of the disciplines by educators. Of importance in future seminars, perhaps more than the specific ideas discussed, will be not only the participants' exhibition of genuine interest and concern for the disciplines and for the curriculum but also the stimulation of a continued and deepened commitment by scholars and educators to explore curriculum problems together. Also, forthcoming project reports may be expected to display the influence of the disciplines as a curriculum source.

While Bruner and Foshay emphasized the importance of the concept of the structure of knowledge for both teaching and learning, neither described the structure of any discipline in detail. Indeed, one of the most penetrating criticisms of their proposals has been the charge of the possible extravagance of their claims. Disagreement is profound on the degree of preciseness with which disciplines may be differentiated. While admittedly insufficient, recent efforts to describe and analyze the structures of disciplines indicated that this task is necessary not only to an understanding of the disciplines but also to the advancement of the cause of using the disciplines directly in the school curriculum.

Drawing on examples from the sciences, Schwab (1962) noted that a structure is comprised of those concepts that define the domain of the discipline and determine its inquiries. Since attention to structure is attention to pattern rather than to catalogue, Schwab argued that Spencer's question, "What knowledge is of most worth?" long of interest to educators (Kazamian, 1960), may well be irrelevant. Johnson's (1962) attempt to describe structure in the social sciences probably did not achieve success because of his search for a common-denominator structure for these related but different disciplines.

English

Parker (1962) admitted that a pedagogical structure in English studies does not exist. He held that the subject of English should consist only of language, literature, and composition. The interrelationships of these elements, while not well recognized, must receive attention if English is to be properly conceptualized, taught, and studied.

In his delightful essay, MacLeish (1961) observed that more agreement may be obtained regarding what English is not than concerning what it is. His suggestion that English ministers between text and "the ground of our being" is poetically reassuring and a reasonable, yet incomplete, guide to practice. However, his affirmation helps set a proper course for serious inquiry into the discipline or area of study that is English; it does not represent a windfall.

Mathematics and Science

Much attention has been focused on the national curriculum development projects in mathematics and the sciences. Most of these projects appear to have been related directly to their disciplines. However numerous the project descriptions that exist, misunderstanding and unintelligent acceptance of "new" curriculums persist. Helpful would be additional examinations, in depth, of the nature of all the disciplines, such as the examination Rees (1962) made for mathematics. Rees asserted that mathematics is basically a language; that mathematical thinking is both inductive and deductive, both intuitive and analytic; that mathematical problems arise both from within and without mathematics; that the discoveries made by mathematics are useful to conceptualizing in other sciences; and that abstraction has produced simplification. Understanding these ideas should be of value in making a variety of decisions about mathematics in the curriculum.

Social Sciences

Efforts to improve social studies instruction, while not so dramatic as those in the natural sciences, have neither neglected nor capitalized on analyses of the nature of the several social sciences. Important strides were taken by a symposium sponsored by the American Council of Learned Societies and National Council for the Social Studies (1962). Much needed, at present, are additional discussions about the nature of the social sciences addressed more to the problem, "What is the discipline of history?" than to the topic, "What understandings in history should teachers have, or what understandings should be employed, in a school history course?"

Progress in defining the essential characteristics of geography and economics for curriculum development purposes was more substantial than was work toward understanding the basic elements of other social sciences. In a publication for teachers associated with the national geography project, the initial chapter was titled "The Nature of Geography" (Pattison, 1962). A high school American history course was suggested by Brown (1960), with a purpose of introducing students to history as a system of ideas and a system of analysis as well as to history as fact. Schrag (1961) reported interesting plans for a history course designed to bring students into closer contact with basic historical materials, methods, and demands. These efforts, which were related to history, like those of Davis (1962), were based on conceptions of the discipline that were more implied than elaborated.

The observation was frequently made that sharp divisions between related disciplines are difficult to conceptualize. Hanna (1962) noted that structure is a human product to be employed as an organizational scheme. From this nominalist position, at such variance with the one sought by Phenix (1962), Hanna proposed that concepts and generalizations drawn from the several social science disciplines form the structural elements of the social studies curriculum. Hanna and Lee (1962) summarized results of a cooperative research project to identify and to group major social science concepts into such categories as "organizing and governing" and "communicating ideas and feelings." This significant research project may be expected to have an observable impact on social studies programs. However, its premise and conclusions may be seriously in error, as viewed from the positions of Phenix (1962) and Schwab (1962), inasmuch as concepts have been isolated from their source of origin and from the methods by which they were discovered and, consequently, by which they may be understood best. Moreover, the efforts may be seen as attempts to restructure knowledge from outside the disciplines. Whatever common elements may be discerned (Brodbeck, 1962) or whatever possible common groupings of disciplines may be formulated (Broudy, 1962; Kimball and McClellan, 1962), to treat the "commonness" aspect as the most important one may be to homogenize the separate disciplines beyond recognition. The issue presented here is fundamental. Its resolution will determine, in no small measure, much of the character of future curriculum planning.

Structures and Groupings of Disciplines

Bellack (1962) and Miel (1963) made notable suggestions for dealing not only with the structures of the disciplines themselves but also with the structures as elements of a comprehensive curriculum design. They developed the possibilities of including both discipline-centered and problem-centered components as advanced by Foshay (1961).

Partly to counteract undue fragmentation of knowledge, Bellack (1962) proposed emphasis on the structures and strategies of groupings of disciplines rather than on the structures and strategies of individual disciplines. Additional work will be necessary, in his rather subdued nominalist suggestion, to distinguish between attention to particular systems of ideas or systems of inquiry and an orientation to a kind of reconceptualized "broad fields" organization.

An integral feature of Bellack's proposal was a coordinating seminar which would provide opportunities for students to deal directly with knowledge bearing on important human affairs. Such a seminar, perhaps similar to a successful venture described by Patterson (1962), would partially meet the expressed hopes of Van Til (1962) and of others who pressed for attention to societal problems in the curriculum.

A significant implication in Bellack's proposal is that, in the different context of studying an important problem, students would put to a second test ideas and modes of inquiry gained from encountering the disciplines. A basic issue involved here, and one of considerable import, relates to the nature of problem solving. Traditionally, problem solving has been interpreted as a general procedure applicable to all problems. On renewed examination, this activity may, with more accuracy and efficiency, be undertaken as one brings to the problem patterns of ideas and inquiry methods specific to particular disciplines.

Miel's (1963) major proposal dealt with disciplines grouped according to families (e.g., arts and natural and physical sciences) and studied comparatively in three cycles in grades 4-12. She suggested how this proposal might operate in both elementary and secondary school divisions and indicated certain problems and issues arising from its operation. Her plan deserves the serious discussion she desired to provoke.

Conclusion

Curriculum theory building, which includes attention to organized knowledge, certainly will continue. For example, treatises on curriculum probably will follow Taba's (1962) lead in including specific attention to the nature of knowledge. To state that the prospects for more adequate curriculum development are possible with the present understanding of the nature of knowledge and of the nature of the disciplines is probably to exhibit immodest naïveté. The necessary substantial research and analysis indicates the obvious emptiness which is confronted by a consuming desire to be filled.

This chapter has reviewed impressive but beginning efforts to analyze organized knowledge and its relationship to the curriculum. Nevertheless, present results may be viewed, for purposes of research, as an inventory of propositions to be investigated. In addition to those problems identified in preceding sections, others deserve mention. One has to do with the balance between the necessary rhetoric of conclusions and the desired process of transaction in learning. Others include the relative importance of learning by discovery and by didactic instruction in specified types of situations; the types of curriculum structures, learning materials, and resources necessary for confronting the disciplines; and the procedures useful in effecting the institutional and attitudinal changes in prospect for the teaching profession and the public.

Continued efforts to describe, in meaningful terms, the nature of the disciplines will be of major consequence. Participation in this activity by the major scholarly societies and prominent scholars as well as by educators is much to be desired. Forums must be provided for responsible analysis and criticism of attempts to translate theoretical formulations into school practice. Needed are schools and teachers open to experimentation employing newly designed curriculum components. In addition to examinations of particular disciplines, attention must also be directed to an analysis of knowledge itself. Helpful in this regard should be the analytic insights provided by Ryle (1949) and illustrated in several penetrating essays edited by Smith and Ennis (1961).

Bibliography

AMERICAN COUNCIL OF LEARNED SOCIETIES and NATIONAL COUNCIL FOR THE SOCIAL STUDIES. *The Social Studies and the Social Sciences.* New York: Harcourt, Brace and World, 1962. 303 pp.

BELLACK, ARNO A. "Conceptions of Knowledge: Their Significance for the Curriculum." *The Nature of Knowledge.* (Edited by William A. Jenkins.) Milwaukee: University of Wisconsin, 1962. pp. 42-52.

BRODBECK, MAY. "Toward a Fabric of Knowledge—Common Elements Among Fields of Learning." *Educational Record* 43: 217-22; July 1962.

BROUDY, HARRY S. "To Regain Educational Leadership." *Studies in Philosophy and Education* 2: 132-58; Spring 1962.

BROWN, W. BURLIE. *United States History: A Bridge to the World of Ideas.* Washington, D.C.: Service Center for Teachers of History, American Historical Association, 1960. 77 pp.

BRUNER, JEROME S. *The Process of Education.* Cambridge, Mass.: Harvard University Press, 1960. 97 pp.

BRUNER, JEROME S. *On Knowing: Essays for the Left Hand.* Cambridge, Mass.: Belknap Press of Harvard University, 1962. 165 pp.

DAVIS, O. L., JR. "A New Look in Materials for Discovering Knowledge in the Social Studies." *High School Journal* 45: 226-34; March 1962.

FOSHAY, ARTHUR W. "A Modest Proposal." *Educational Leadership* 18: 506-16, 528; May 1961.

HANNA, PAUL R. "Structure of Knowledge: The Interrelationship of Ideas." *The Nature of Knowledge.* (Edited by William A. Jenkins.) Milwaukee: University of Wisconsin, 1962. pp. 68-82.

HANNA, PAUL R., and LEE, JOHN R. "Generalizations from the Social Sciences." *Social Studies in Elementary Schools.* (Edited by John V. Michaelis.) Thirty-Second Yearbook. Washington, D.C.: National Council for the Social Studies, a department of the National Education Association, 1962. pp. 62-89.

JOHNSON, EARL S. "The Concept of Structure in the Social Sciences." *Educational Record* 43: 206-209; July 1962.

KAZAMIAN, ANDREAS M. "'What Knowledge Is of Most Worth?' An Historical Conception and a Modern Sequel." *Harvard Educational Review* 30: 307-30; Fall 1960.

KIMBALL, SOLON T., and MCCLELLAN, JAMES E., JR. *Education and the New America.* New York: Random House, 1962. Chapter 13, "Education for Commitment," pp. 279-304.

MACLEISH, ARCHIBALD. "What Is English?" *Saturday Review* 44: 12-14; December 9, 1961.

MIEL, ALICE. "Knowledge and the Curriculum." *New Insights and the Curriculum.* 1963 Yearbook. Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1963. Chapter 4, pp. 71-104.

NATIONAL EDUCATION ASSOCIATION, PROJECT ON INSTRUCTION. *The Scholars Look at the Schools: A Report of the Disciplines Seminar.* Washington, D.C.: the Association, 1962. 64 pp.

PARKER, WILLIAM R. "The Concept of Structure in English." *Educational Record* 43: 210-16; July 1962.

PATTERSON, FRANKLIN. *Public Affairs and the High School.* Medford, Mass.: Lincoln Filene Center for Citizenship and Public Affairs, Tufts University, 1962. 43 pp.

PATTISON, WILLIAM D., project director. *Advisory Paper for Teachers Associated with the High School Geography Project.* Los Angeles: Joint Committee on Education of the Association of American Geographers and the National Council for Geographic Education (Department of Geography, University of California), 1962. 73 pp.

PHENIX, PHILIP H. "The Disciplines as Curriculum Content." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 57-65.

REES, MINA. "The Nature of Mathematics." *Science* 138: 9-12; October 5, 1962.

RYLE, GILBERT. *The Concept of Mind.* London: Hutchinson and Co., 1949. 334 pp.

SCHRAG, PETER. "Joint Action in Ivied Halls." *Saturday Review* 44: 54-55, 65; October 21, 1961.

SCHWAB, JOSEPH J. "The Concept of the Structure of a Discipline." *Educational Record* 43: 197-205; July 1962.

SMITH, B. OTHANEL, and ENNIS, ROBERT H., editors. *Language and Concepts in Education.* Chicago: Rand McNally & Co., 1961. 221 pp.

TABA, HILDA. *Curriculum Development: Theory and Practice.* New York: Harcourt, Brace and World, 1962. 529 pp.

VAN TIL, WILLIAM. "What Are the Sources of Curriculum?" *What Are the Sources of the Curriculum? A Symposium.* Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1962. pp. 35-47.

Additional References

AUSUBEL, DAVID P. "In Defense of Verbal Learning." *Educational Theory* 11: 15-25; January 1961.

DAVIS, O. L., JR. "The Teaching of the Disciplines and the Core Program: Some Comments and Questions." *High School Journal* 44: 239-47; April 1961.

FOSHAY, ARTHUR W. "What Is the Message?" *Saturday Review* 43: 35, 58-61; February 13, 1960.

FOSHAY, ARTHUR W. "Three Illustrations of Excellence." *Elementary School Journal* 62: 77-81; November 1961.

FOSHAY, ARTHUR W. "Discipline-Centered Curriculum." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 66-71. (a)

FOSHAY, ARTHUR W. "Knowledge and Structure of the Disciplines." *The Nature of Knowledge.* (Edited by Willian A. Jenkins.) Milwaukee: University of Wisconsin, 1962. pp. 28-40. (b)

FRASER, DOROTHY M. *Current Curriculum Studies in Academic Subjects.* Washington, D.C.: Project on Instruction, National Education Association, 1962. 102 pp.

HERRICK, VIRGIL E. "Sources of Curriculum Development." *What Are the Sources of the Curriculum? A Symposium.* Washington, D.C.: Association for Supervision and

Curriculum Development, a department of the National Education Association, 1962. pp. 60-71.

HUGHES, PHILLIP. "Decisions and Curriculum Design." *Educational Theory* 12: 187-92; July 1962.

JAROLIMEK, JOHN. "Curriculum Content and the Child in the Elementary School." *Social Education* 26: 58-62, 117-20; February 1962.

JOHNSON, EARL S. "Ways of Knowing." *The Nature of Knowledge.* (Edited by William A. Jenkins.) Milwaukee: University of Wisconsin, 1962. pp. 16-26.

MCCLELLAN, JAMES E. "The Logical and the Psychological: An Untenable Dualism?" *Language and Concepts in Education.* (Edited by B. Othanel Smith and Robert H. Ennis.) Chicago: Rand McNally & Co., 1961. Chapter 10, pp. 144-60.

NATIONAL EDUCATION ASSOCIATION, PROJECT ON INSTRUCTION. *The Principals Look at the Schools: A Status Study of Selected Instructional Practices.* Washington, D.C.: the Association, 1962. 75 pp.

PHENIX, PHILIP H. "Key Concepts and the Crisis in Learning." *Teachers College Record* 58: 137-43; December 1956.

ROGERS, VINCENT R. "History for the Elementary School Child." *Phi Delta Kappan* 44: 132-35; December 1962.

SAND, OLE, AND MILLER, RICHARD I. "Perspective on National Studies in the Disciplines." *Journal of Secondary Education* 38: 27-33; January 1963.

SUTHERLAND, NEIL. "Structure in the History Curriculum." *Social Education* 26: 133-36; March 1962.

WILHELMS, FRED T. "Curriculum Sources." *What Are the Sources of the Curriculum? A Symposium.* Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1962. pp. 14-25.

CHAPTER III

Societal Forces Influencing Curriculum Decisions

GORDON F. VARS and WILLIAM T. LOWE

THIS CHAPTER deals with societal forces that directly influence the selection of subject matter or learning experiences to be provided by the schools. Excluded are developments within a discipline that may have changed the nature of content in a field, but that did not have explicit implications for curriculum planning. Also excluded are most studies generally dealing with social forces affecting the school as a total institution. A few particularly significant items of the latter are listed as additional references. Important findings in some studies are omitted because they were not specifically pertinent to curriculum decision making; the reader should not infer that the conclusions summarized here exhaust the significance of any particular study.

Social realities such as race, social class, peer culture, and the family are dealt with first, followed by social problems such as school dropouts, delinquency, and politics. Since society is an entity, it is difficult to establish categories; many studies deal with interrelationships among the various aspects of society. Analyzing forces in a living, changing, incredibly complex society while living within it is no easier for these reviewers than it was for their predecessors, Breslow and others, in the June 1960 REVIEW. In some cases, the writers have supplied their own interpretations of a study's curriculum implications. Inevitably, this chapter reflects the backgrounds and biases of the reviewers, both in the studies selected and in the implications inferred. The reader is urged to examine research in the original and to draw his own conclusions.

Social Realities

Racial and Ethnic Differences

Recent research in intergroup relations appears to add little to the generalizations of Van Til in the October 1959 REVIEW. One of the few studies to draw curriculum implications from the desegregation process was Wiley's (1961) study of desegregation in Evansville, Indiana. Here, the smooth transition was attributed largely to factors in the community, such as the high respect accorded to the superintendent and members of the board of education. Even after desegregation had taken place, there was no planned program for discovering intergroup relations problems in the schools. Moreover, only a very small percentage of the staff thought that human relations problems were important enough to warrant the refocusing of educational objectives or the reconsidering of school practices. In spite of the Evansville study, widespread repercussions of racial conflict argue that a school facing desegregation should develop and maintain the best possible human relations program.

Attitudes in this area are difficult to change, of course. Campbell (1961) employed questionnaires, interviews, role playing, and a sophisticated social-distance scale to measure prejudice in a Southern university. Testing the hypothesis that those who support racial segregation suffer from moral guilt because of the inconsistency of these tenets with the "American creed," he found very little guilt—students either did not believe in the American creed or felt that the best way to follow the creed was to "keep the Negro in his place." Teaching about democratic ideals and values will not, in itself, change attitudes about segregation.

This is not to say that racial and ethnic attitudes cannot be changed. From his examination of existing research, Yoshino (1961) concluded that the teacher, curriculum, and experiences with fellow students can influence these feelings, and that deliberate effort at improving ethnic understanding does help young elementary school children. Scott's (1961) study of work groups demonstrated that when management places together adults of equal ability but of differing racial and ethnic groups to complete a job, human relationships improve. Key factors seem to be the equal status of members and the task orientation of the group, both of which should be kept in mind when using the group-work approach in schools to reduce intergroup tensions.

Additional References: Hodgkinson (1962); Tumin (1957); U.S. Commission on Civil Rights (1961; 1962a, b).

Religion

The interaction between organized religion and the public schools continued to be a source of conflict. Political and ethnic issues frequently confounded debate over the place of religion in the curriculum. Bible reading, prayers, religious holiday observances, released time, transportation of parochial school students at public expense, exemption from health instruction on religious grounds—all these were involved in case studies of local conflict in New Jersey, which were examined by Baker (1961). Extensive interviewing and study of local press dispatches led Baker to conclude that population shifts played an important role in the degree of conflict. As Catholics and Jews moved out of the cities into the suburbs and outlying areas, they challenged traditional Protestant domination in many communities. Since this dominance was often reflected most dramatically in the public schools, it was there that the conflict occurred. Also important was what Baker termed "idiosyncratic individual behavior," actions of community members seemingly motivated entirely by personal experiences and biases. There was no evidence that outside organizations, such as civil liberties groups, played any significant part in initiating conflict. The degree of conflict was affected by the nature of news coverage, actions of the superintendent, organization of protesting groups, and policies of the state education department. Clearly, local community factors must be taken into account in dealing with religion in the curriculum.

Similar conclusions were drawn by Boles (1961), who made a careful and systematic review of legislation, court decisions, and traditions affecting church relations with the public schools. He called for an objective study of religion as one important aspect in the development of man. He warned, however, that any practice that attacks or supports a particular sectarian point of view will cause increasing trouble as the population becomes more mobile and communities become less isolated from society-at-large.

Strong religious values tend to be associated with high ratings of dogmatism, according to Frumkin's (1961) research with college students. Examining the relationships among dogmatism, social class, values, and academic achievement in sociology, he found that, regardless of other variables, the more education the student had, the less dogmatic he was. Since other studies have contended that dogmatism, or authoritarianism, is related to social, economic, ethnic, and racial prejudice, Frumkin concluded that instruction that reduces dogmatism in some areas tends to open students' minds in others.

If Frumkin is correct, a curriculum designed to reduce prejudice in any area is likely to affect students' religious values, with possible repercussions from parents and church authorities. Rosenblum's (1959) research suggests that these reactions will differ according to the degree of orthodoxy of the religious group. He found that in religious groups classed as liberal, the more religious the person, the less his ethnic prejudice; with orthodox groups, the relationship was reversed. Since public schools are committed to broadening students' horizons, they must be aware of students' present values; offer effective personal guidance for students trying to solve value conflicts; and keep channels of communication open among home, school, and church, so that conflicts can be minimized.

Additional References: Corcoran (1960); Lenski (1961); Travers and Davis (1961); Tumin (1961).

Socioeconomic Class

The pervading influence of social class received additional documentation by Havighurst and others (1962) in *Growing Up in River City*. Approximately 500 young people from age 11 to 20 were studied as they grew up in a typical Midwestern city with a population of 45,000. Through a variety of measures, marked relationships were found between socioeconomic status and intelligence, adjustment, achievement, educational motivation, educational level attained, delinquency, church activity and interest, age of first marriage, work adjustment, and initial adult adjustment. Many of these factors interacted. Considerable social mobility was noted, with educational achievement the usual way to rise in status. The school was urged to make up for the disadvantages with which some children start out in life and to provide a climate which minimizes conflicts between socioeconomic classes.

Comparing mill (lower class) and town (higher class) sections of a Southern community, Morland (1960) found significantly lower educational aspirations among lower class children, but no significantly different occupational aspirations. This study may suggest that guidance and instruction should stress the relationship between education and career advancement.

Rutledge (1960) also called for improved guidance and for adjustment of the curriculum to meet the needs of different status groups. He found a number of significant relationships between social status and the personal-social problems of young adolescents revealed by questionnaires and by the *Mooney Problem Check List*. Though sex was not a significant factor in the number of problems checked, social status apparently had more impact on the girls than on the boys. As might be expected, the incidence of problems was highest for students of low social status, lowest for those of high status. When comparing achievement of Spanish American and Anglo-American children, Cline (1961) found that biculturalism by itself does not affect achievement, but that socioeconomic class does.

Either explicitly or implicitly, these studies seem to suggest special curriculum provisions for children from the lower socioeconomic levels. This is truly an explosive issue if it is taken to mean conscious segregation of children by socioeconomic class. Conventional ability grouping is attacked even now on the grounds that it segregates social classes. Hence, it is difficult to see how such a proposal could win widespread acceptance. On the other hand, guidance-centered programs carried out in heterogeneous groups might provide needed special attention without labeling students.

Additional References: Dawson (1961); Fisher (1961); King (1961); Robins, Gyman, and O'Neal (1962).

Peer Culture

The nature and importance of the adolescent subculture was further documented by Coleman (1961) in his analysis of the social systems of 10 high schools in widely varied communities. Stressing differences between adolescent values and those of adult society, he contrasted the passive dependency which the school imposes upon adolescents with the social emancipation granted them by their parents. Coleman asserted that intellectual activities must capture the excitement of doing, exploring, creating, and meeting a challenge, if they are to compete with athletic, social, and extracurricular activities for students' energy. Even more significant were the implications for the cocurriculum, since interscholastic athletics was seen as a major contributor to the status achieved by high school boys, and popularity or social success, the major status factors for girls. To channel adolescent energy into academic pursuits, Coleman proposes interschool competition in scholastic matters, such as intellectual games and problems.

Additional Reference: Smith (1962).

Family

As the family changes, so must those aspects of the curriculum devoted to family life education. Studies such as those by Roy (1961) and Weil (1961) indicated that employment of both parents is not necessarily harmful to their children in terms of academic achievement, social life, or proneness to delinquency. These studies suggested, among other things, that both course content and guidance programs should help girls to consider the circumstances under which a career and homemaking may be combined successfully.

Havighurst and others (1962) found many relationships between early marriage, low intelligence, low social class status, and school dropouts. Furthermore, nearly all the young couples studied rated low in marital adjustment. Nevertheless, they concluded that early marriage might be the most desirable course for girls who were unsuccessful in school; they recommended a special curriculum for them, organized around their roles as wives and mothers-to-be. The teacher-counselor for the girls would be a combination social worker, home economics teacher, and club adviser. Activities for the older girls, which also would be available to those who had already left school, would be designed to give group members a source of needed emotional support as they made the transition to married life.

Turner (1961), on the other hand, was more concerned about reducing the incidence of student marriages. From his survey of student marriages in Missouri high schools, he concluded that one of the most effective deterrents to student marriages is a challenging secondary school program, operating in cooperation with all youth-serving agencies in the community.

Additional References: Brim, Fairchild, and Borgatta (1961); Dentler and Hutchinson (1961); Dentler and Monroe (1961); Riessman (1962); Westoff and others (1961).

Social Problems

The social realities described earlier interact in numerous ways to produce pressing social problems, particularly in the nation's large cities.

Dropouts

Though not a research report in the usual sense, Conant's (1961) *Slums and Suburbs* called the nation's attention to the "social dynamite" accumulating in the big cities as the result of dropouts and unemployment and complicated by social class and intercultural problems. Contrasting educational opportunities in slums and suburbs, he called for improved guidance services and for realistic vocational preparation, including work-study programs, for city youth. Burchill (1962) cited numerous detailed case studies to support the value of work-study programs for disadvantaged or "alienated" youth.

New York's School to Employment Program (STEP) and Chicago's Educational and Employment (Double EE) experiment exemplify existing special curriculum arrangements for dropouts and potential dropouts. Unfortunately, in a study completed in 1961, Rioux (1961) found that less than one-fifth of the 100 largest cities had any special provisions for dropouts and that only a small proportion of the others had plans to develop such programs. Most special provisions were primarily vocational or prevocational and directed toward preparing the student for employment in a short space of time. Havighurst and others (1962) pointed out that a community-wide effort is required to solve the dropout problem. Besides work experience for boys, they advocated the preparation for marriage described earlier in this chapter. The study by Lichter and others (1962) of intellectually able potential dropouts showed that no one plan works best with all youngsters.

Like Havighurst and his associates, Lichter and others found that adequate success experience in school was a key factor in a student's continuation. Johnson (1960), too, noted that two-thirds of the dropouts he interviewed had been school failures. In spite of this, the dropouts as a group did not dislike either school or the teachers and probably would have completed school if they had been encouraged.

The Maryland survey reported by Williams (1963) also indicated that most dropouts were fairly normal youths, 49.8 percent being average or above average in intelligence, 70 percent living with both parents, and 79 percent not considered serious behavior problems by either their counselors or their parents. Although he concluded that major influences causing early withdrawal came from outside the school, Johnson (1960) called for a number of curriculum revisions to reduce dropouts. He recommended that participation in extraclass activities be encouraged and guidance services be extended, with special individualized attention to those with low grades and irregular attendance. The entire staff, he concluded, should concentrate on motivating students through meaningful course content and through interesting instruction.

More drastic curriculum modifications were advocated by Chandler, Stiles, and Kitsuse (1962) in an interdisciplinary symposium. Besides spending more time in the study of social problems, schools should aid desirable social changes like urban renewal by deliberately attempting to re-educate relocated slum dwellers. Advocating drastic overhaul of metropolitan school organization and administration, Havighurst (1961) recommended that special attention be given to the culturally deprived at the elementary level and at work-study centers for children who cannot profit from regular schools. New York City's Higher Horizons project has demonstrated the value of cultural enrichment for disadvantaged children.

It is not just the large cities that are in difficulty because of demographic changes. In *Plainville Fifteen Years Later*, Gallaher (1961) documented the woes of a small rural town: economic decline, paucity of cultural advantages, rigid class system, and exodus of capable youth. Curriculum implications that are left to the reader would likely include attention to community problems in the school, an improved school cultural program, and vocational preparation geared to a mobile student population.

Additional References: National Education Association and American Association of School Administrators, Educational Policies Commission (1962); U.S. Department of Health, Education, and Welfare, Office of Education (1962).

Delinquency

Delinquency, too, seems to be related to success in school. Havighurst and others (1962) found delinquency among River City youth to be concentrated among those of lower socioeconomic class who had dropped out of school, had relatively low intelligence, and came from broken or inadequate homes. They postulated that at least some students turn to delinquency because they are unable to achieve success according to the usual standards prevailing in the community. Elliott's (1962) questionnaire study tended to confirm this point of view. He found that both delinquents and nondelinquents held similar views of the role played by a given occupation or level of education in achieving success, but delinquents, regardless of their social class, perceived themselves as having had less opportunity for achieving success.

Karacki and Toby (1962) investigated a particular juvenile gang to test the commonly held assumption that this form of delinquent behavior stems from economic and social deprivation. On the contrary, members of the gang were all white and nearly all of the same religious group, middle class, and from privately owned homes. The authors concluded that failure to develop commitment to adult roles and adult values was the basic source of gang cohesiveness. Gang members held to the values of the peer culture, such as demand for immediate gratification of their desires, loyalty to peers, and assertion of masculinity through physical aggression. Karacki and Toby concluded that delinquency could be reduced, even in blight areas, by developing, through schools and other social institutions, opportunities for the adolescent to develop allegiance to the values prevalent in adult culture.

Additional References: Kinch (1962); Martin and others (1961).

Political Problems

Though cold war tensions must inevitably influence curriculum decisions, there appears to be little empirical research to document this fact. Feldman (1960) queried city school administrators about instruction concerning the United Nations Educational, Scientific, and Cultural Organization. He found little instructional use of UNESCO materials and concluded that controversy over UNESCO in the Los Angeles area might have further reduced the study of the agency in that section of California.

Local, state, and national politics also influence curriculum decisions. After studying political forces affecting school finances in eight Northeastern states, Bailey and others (1962) concluded that it is essential for educators to become active in politics. Robinson (1961) came to a similar conclusion from his review of selected legislation on school matters. He found rigid, conflicting legislation in many regions of the nation and concluded that one of the best ways to avoid curriculum rigidity is to provide federal funds to the states for broad educational purposes. Each state would have wide discretion and freedom to decide how the funds would be allocated to local districts.

Both Jordan (1961) and Samalonis (1962) studied the influence of state legislatures upon curriculum. Jordan concluded that Indiana statutes contain an excessive number of mandatory requirements which tend to make the Department of Public Instruction a regulatory agency rather than a source of leadership. After examining existing laws in all 50 states, Samalonis stated that the most common practice of legislatures was to make subject matter prescriptions without explicitly stating objectives or clarifying the intent of the statute. Physical fitness and the social sciences were the most highly regulated subject matter areas, in terms of both time allotment and grade placement.

Pressure groups also make their influences felt upon the curriculum. Farnand (1959) studied the social philosophies of three prominent groups opposed to federal aid to education: the American Farm Bureau Federation, the Chamber of Commerce of the United States, and the National Association of Manufacturers. He characterized all three as conservative, materialistic, and pragmatic. Decentralized government was favored by these groups, who believe that government should be primarily political, not economic, but that it should protect business. Since these organizations furnish th- schools with a great deal of free and inexpensive instructional material, their points of view are likely to be reflected in the curriculum. This is not to suggest that biased materials be eliminated, but that teachers must use them with discretion and must be aware of the points of view represented.

Selakovich (1962) studied techniques used by pressure groups attempting to influence the teaching of American history; namely, the American Legion, the Daughters of the American Revolution, the Conference of American Small Business Organizations, the Friends of the Public Schools, and the National Economic Council. These groups held similar concepts of Americanism, favored a strict interpretation of the U.S. Constitution, and endorsed positions generally hostile toward the United Nations. Charges of disloyalty or of subversion were leveled by these groups against both teachers and the instructional materials used in schools. Teachers also were accused of being ignorant of the Constitution and of failing to teach Americanism. On the other hand, Kitaguchi (1962) found that in Hawaii, at least, the points of view expressed by organized labor groups did not differ significantly from those of a panel of educators.

Although numerous groups seem ready to pounce upon the schools if they deal with controversial issues, some studies indicate that there may be less community opposition than is commonly assumed. Such was the conclusion of Coan's (1961) study of parents and teachers in Kansas. Rosenthal (1961) found in his study of two communities in New Jersey that educators generally had a rather wide scope for independent action. This was particularly true on narrow audience issues such as curriculum and personnel, as opposed to wide audience issues having to do with budgets and bond issues. A complicating element mentioned by Cooper (1959) was the fact that teachers, principals, and superintendents interpret external pressures differently. Sometimes, they do not even recognize the forces being exerted upon their curriculum and teaching.

Additional References: Ayars (1961); Martin (1962); Munger and Fenno (1962); Newman (1961); Palmer (1961); Rossi and Dentler (1961); Schmidt (1962); Todd (1962).

Economic Influences

Continued prosperity and booming technology characterize the economic situation in which curriculum decisions are being made today. Neither of these appears to have been extensively researched. Rhoades (1961) found wide variations in the degree of optimism reflected in publications dealing with the implications of automation for education. More advanced technical training seemed to be indicated, although automation might require less training in specific skills and more attention to basic sciences and general education. Students should be prepared for change, and their education should be oriented more to the future than it is at present. Two likely outgrowths of automation—technological unemployment and a shorter workweek—highlight the need for education in the wise use of leisure time. Ford's (1962) analysis of leisure time activities of aged residents of Indiana indicated that schools of the past could take little credit for interesting these individuals in satisfying recreational activities.

Educators continued to study both the effects of technology upon students and ways to introduce better technology into the schools. Schramm, Lyle, and Parker (1961) studied television's impact upon children in a large study involving 6,000 children and 2,300 teachers and parents. They concluded that, for most children under most conditions, television was neither particularly beneficial nor very harmful. One fact of importance to curriculum planners is that children with average intelligence seem to have wider vocabularies if they are frequent viewers. Previous studies have suggested that television may be sending children to school with a broad but superficial acquaintance with much of the world that they are too young to understand. Further research on this question, particularly at the local level, should enable schools to adapt the curriculum to the actual sophistication of the student population.

Of deeper significance, perhaps, is the image of American character portrayed by this influential medium. After a year's study of network television programs, Bell (1961) characterized this image as that of a middle class, middle-aged, middle-of-the-road male—a nonheroic figure in a nonheroic society. His interest is in immediate, personal problems, not in philosophy or in world problems, and he is governed by relative values. Schools must deal with this image as they attempt to modify student values through curricular experiences.

Conclusions

Examination of research on societal forces affecting curriculum decisions leads the writers to the following tentative conclusions:

1. The basic assumption of the influences of societal forces upon decisions concerning the curriculum is supported by the research reviewed. What is taught, why it is taught, and how it is taught are affected by dynamic social forces. Study of these forces is a continuing responsibility of all educators.
2. There is a growing body of basic and applied scholarship in this area. The number of relevant doctoral studies is encouraging. Because of the nature of the topic, this chapter has emphasized applied research.
3. There still is not enough research in this field. Significant questions remain almost totally unstudied, especially the curricular implications of some of the basic social problems and changes.
4. A number of studies simply support or replicate older research evidence that schools all too frequently do not apply research findings already available.
5. Numerous studies emphasize the importance of providing curricular experiences that are meaningful to the learner in terms of the cultural context in which he is living.

The additional references listed below represent more comprehensive and general overviews of the area than do most of the sources that have been considered in this chapter.

Additional References: Bell (1962); Bienenstok and Sayres (1962); Elkin (1960); Gross, Wronski, and Hanson (1962); Halsey, Floud, and Anderson (1961); Ladd and Sayres (1962); National Society for the Study of Education (1961).

Bibliography

Bailey, Stephen K., and others. *Schoolmen and Politics: A Study of State Aid to Education in the Northeast.* Syracuse, N.Y.: Syracuse University Press, 1962. 111 pp.

Baker, John Henry. *Church and State in New Jersey Public Education.* Doctor's thesis. Princeton, N.J.: Princeton University, 1961. 292 pp. Abstract: *Dissertation Abstracts* 23: 284-85; No. 1, 1962.

Bell, Richard Henry. *A Study of the Image of the American Character as Presented in Selected Network Television Dramas.* Doctor's thesis. Columbus: Ohio State University, 1961. 192 pp. Abstract: *Dissertation Abstracts* 22: 2242-43; No. 7, 1962.

Boles, Donald E. *The Bible, Religion, and the Public Schools.* Ames: Iowa State University Press, 1961. 308 pp.

Burchill, George W. *Work-Study Programs for Alienated Youth: A Casebook.* Chicago: Science Research Associates, 1962. 265 pp.

Campbell, Ernest Q. "Moral Discomfort and Racial Segregation—An Examination of the Myrdal Hypothesis." *Social Forces* 39: 228-34; March 1961.

Chandler, B. J.; Stiles, Lindley J.; and Kitsuse, John I., editors. *Education in Urban Society.* New York: Dodd, Mead, and Co., 1962. 279 pp.

Cline, Marion, Jr. *Achievement of Bilinguals in Seventh Grade by Socioeconomic Levels.* Doctor's thesis. Los Angeles: University of Southern California, 1961. 195 pp. Abstract: *Dissertation Abstracts* 22: 3113-14; No. 9, 1962.

Coan, Clark. *A Study of the Attitude of Selected Social Studies Teachers and Parents of Kansas High School Students Regarding the Inclusion of Controversial Issues as Part of the Secondary School Social Studies Program.* Doctor's thesis. Lawrence: University of Kansas, 1961. 317 pp. Abstract: *Dissertation Abstracts* 22: 2246; No. 7, 1962.

Coleman, James S. *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education.* New York: Free Press of Glencoe, 1961. 368 pp.

Conant, James B. *Slums and Suburbs: A Commentary on Schools in Metropolitan Areas.* New York: McGraw-Hill Book Co., 1961. 147 pp.

Cooper, John Ellsworth. *An Analysis of Certain External Pressures Affecting Teaching and Curriculum in Selected Elementary Schools of Indiana.* Doctor's thesis. Bloomington: Indiana University, 1959. 327 pp. Abstract: *Dissertation Abstracts* 20: 3590-91; No. 9, 1960.

Elliott, Delbert S. "Delinquency and Perceived Opportunity." *Sociological Inquiry* 32: 216-27; Spring 1962.

Farnand, Everest John Michael, O.F.M. *A Study of the Social Philosophies of Three Major Interest Groups Opposed to Federal Aid to Education.* Doctor's thesis. St. Louis, Mo.: St. Louis University, 1959. 152 pp. Abstract: *Dissertation Abstracts* 20: 4310; No. 11, 1960.

Feldman, Edwin. *UNESCO and American Schools.* Doctor's thesis. Palo Alto, Calif.: Stanford University, 1960. 313 pp. Abstract: *Dissertation Abstracts* 21: 2942-43; No. 10, 1961.

Ford, Phyllis Marjorie. *An Analysis of Leisure Time Activities and Interests of Aged Residents of Indiana.* Doctor's thesis. Bloomington: Indiana University, 1962. 307 pp. Abstract: *Dissertation Abstracts* 22: 3763; No. 10, 1962.

Frumkin, Robert M. "Dogmatism, Social Class, Values, and Academic Achievement in Sociology." *Journal of Educational Sociology* 34: 398-403; May 1961.

Gallaher, Art, Jr. *Plainville Fifteen Years Later.* New York: Columbia University Press, 1961. 301 pp.

Havighurst, Robert J. "Metropolitan Development and the Educational System." *School Review* 69: 251-67; Autumn 1961.

Havighurst, Robert J., and others. *Growing Up in River City.* New York: John Wiley & Sons, 1962. 189 pp.

Johnson, Harvey Clayton. *Pupil Dropouts in the High Schools of Beaumont, Texas, and Curricular Implications.* Doctor's thesis. Los Angeles: University of Southern California, 1960. 311 pp. Abstract: *Dissertation Abstracts* 21: 519; No. 3, 1960.

Jordan, Kenneth Forbis. *An Analysis and Evaluation of Legislative Policy Concerning Public School Curriculum in Indiana.* Doctor's thesis. Bloomington: Indiana University, 1961. 163 pp. Abstract: *Dissertation Abstracts* 22: 779; No. 3, 1961.

Karacki, Larry, and Toby, Jackson. "The Uncommitted Adolescent: Candidate for Gang Socialization." *Sociological Inquiry* 32: 203-15; Spring 1962.

Kitaguchi, Lillian Saeko. *Organized Labor and Public Education in Hawaii, 1945-60.* Doctor's thesis. New York: New York University, 1962. 521 pp. Abstract: *Dissertation Abstracts* 23: 1995-96; No. 6, 1962.

Lichter, Solomon O., and others. *The Dropouts: A Treatment Study of Intellectually Able Students Who Drop Out of High School.* New York: Free Press of Glencoe, 1962. 302 pp.

Morland, J. Kenneth. "Educational and Occupational Aspirations of Mill and Town School Children in a Southern Community." *Social Forces* 39: 169-75; December 1960.

Rhoades, Roger Hanson. *Implications of Automation for Education: A Study of Expectations.* Doctor's thesis. Boulder: University of Colorado, 1961. 153 pp. Abstract: *Dissertation Abstracts* 22: 3474-75; No. 10, 1962.

Rioux, Joseph William. *The Extent, Form, and Future of Educational Provisions for Dropouts.* Doctor's thesis. Detroit: Wayne State University, 1961. 304 pp. Abstract: *Dissertation Abstracts* 22: 3553-54; No. 10, 1962.

Robinson, Glen. "Legislation Influences Curriculum Development." *Educational Leadership* 19: 26-30; October 1961.

Rosenblum, Abraham L. *Social Class Membership and Ethnic Prejudice in Cedar City.* Doctor's thesis. Los Angeles: University of Southern California, 1959. 276 pp. Abstract: *Dissertation Abstracts* 20: 4749; No. 12, 1960.

Rosenthal, Alan. *Community Leadership and Public School Politics: Two Case Studies.* Doctor's thesis. Princeton, N.J.: Princeton University, 1961. 526 pp. Abstract: *Dissertation Abstracts* 22: 2053; No. 6, 1961.

Roy, Prodipto. "Maternal Employment and Adolescent Roles: Rural-Urban Differentials." *Marriage and Family Living* 23: 340-49; November 1961.

Rutledge, Charles Walter. *Social Status and the Perception of Personal-Social Problems of a Selected Group of Early Adolescents.* Doctor's thesis. East Lansing: Michigan State University, 1960. 120 pp. Abstract: *Dissertation Abstracts* 21: 1112-13; No. 5, 1960.

Samalonis, Bernice Lucy. *An Analysis of the Role of the State Legislature in Prescribing the Curriculum.* Doctor's thesis. Urbana: University of Illinois, 1962. 153 pp. Abstract: *Dissertation Abstracts* 23: 499; No. 2, 1962.

Schramm, Wilbur; Lyle, Jack; and Parker, Edwin B. *Television in the Lives of Our Children.* Stanford, Calif.: Stanford University Press, 1961. 324 pp.

Scott, Woodrow W. "The Balanced Ethnic Work Group." *Sociology and Social Research* 45: 196-201; January 1961.

Selakovich, Dan. *The Techniques of Certain Pressure Groups Attempting To Influence the Teaching of American History and Government in the Secondary School 1945-1960.* Doctor's thesis. Boulder: University of Colorado, 1962. 420 pp. Abstract: *Dissertation Abstracts* 23: 1632-33; No. 5, 1962.

Turner, James William. *A Survey of Student Marriage in Missouri Public High Schools.* Doctor's thesis. Columbia: University of Missouri, 1961. 336 pp. Abstract: *Dissertation Abstracts* 22: 2664; No. 8, 1962.

Weil, Mildred W. "An Analysis of the Factors Influencing Married Women's Actual or Planned Work Participation." *American Sociological Review* 26: 91-96; February 1961.

Wiley, Alfred D. *A Study of Desegregation in the Evansville School Corporation, Evansville, Indiana.* Doctor's thesis. Bloomington: Indiana University, 1961. 302 pp. Abstract: *Dissertation Abstracts* 22: 3477; No. 10, 1962.

Williams, Percy V. "School Dropouts." *NEA Journal* 52: 10-12; February 1963.

Yoshino, I. Roger. "Children, Teachers and Ethnic Discrimination." *Journal of Educational Sociology* 31: 391-97; May 1961.

Additional References

Ayars, Albert L. "Role of Interest Groups in Curriculum Planning." *Educational Leadership* 19: 35-38; October 1961.

Bell, Robert R., editor. *Sociology of Education: A Sourcebook.* Homewood, Ill.: Dorsey Press, 1962. 368 pp.

Bienenstok, Theodore, and Sayres, William C. *Contributions of Sociology and Anthropology to Education.* Albany: New York State Education Department, 1962. 61 pp.

Brim, Orville G.; Fairchild, Roy W.; and Borgatta, Edgar F. "Relations Between Family Problems." *Marriage and Family Living* 23: 219-26; August 1961.

Corcoran, Gertrude Beatty. *Social Relationships of Elementary School Children and the Released-Time Religious Education Program.* Doctor's thesis. Palo Alto, Calif.: Stanford University, 1960. 126 pp. Abstract: *Dissertation Abstracts* 21: 2603-604; No. 9, 1961.

Dawson, George G. "Doctoral Studies on the Relationship Between the Labor Movement and Public Education." *Journal of Educational Sociology* 34: 260-69; February 1961.

Dentler, Robert A., and Hutchinson, J. Glenn. "Socio-Economic Versus Family Membership Status as Sources of Family Attitude Consensus." *Child Development* 32: 249-54; June 1961.

Dentler, Robert A., and Monroe, Lawrence J. "The Family and Early Adolescent Conformity and Deviance." *Marriage and Family Living* 23: 241-47; August 1961.

Elkin, Frederick. *The Child and Society: The Process of Socialization.* New York: Random House, 1960. 121 pp.

Fisher, Robert J. "Who Is This Lower-Class Child?" *Journal of Educational Sociology* 34: 309-11; March 1961.

Gross, Carl H.; Wronski, Stanley P.; and Hanson, John W. *School and Society: Readings in the Social and Philosophical Foundations of Education.* Boston: D. C. Heath and Co., 1962. 666 pp.

Halsey, A. H.; Floud, Jean; and Anderson, C. Arnold. *Education, Economy, and Society: A Reader in the Sociology of Education.* New York: Free Press of Glencoe, 1961. 625 pp.

Hodgkinson, Harold L. *Education in Social and Cultural Perspectives.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 243 pp.

Kinch, John W. "Self-Conceptions of Types of Delinquents." *Sociological Inquiry* 32: 228-34; Spring 1962.

King, Morton B. "Socioeconomic Status and Sociometric Choice." *Social Forces* 39: 199-206; March 1961.

Ladd, Edward T., and Sayres, William C., editors. *Social Aspects of Education: A Casebook.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 388 pp.

Lenski, Gerhard. *The Religious Factor: A Sociological Study of Religion's Impact on Politics, Economics, and Family Life.* Garden City, N.Y.: Doubleday & Co., 1961. 381 pp.

Martin, Roscoe C. *Government and the Suburban School.* Syracuse, N.Y.: Syracuse University Press, 1962. 115 pp.

Martin, Roscoe C., and others. *Decisions in Syracuse.* Bloomington: Indiana University Press, 1961. 368 pp.

Munger, Frank J., and Fenno, Richard F., Jr. *National Politics and Federal Aid to Education.* Syracuse, N.Y.: Syracuse University Press, 1962. 193 pp.

National Education Association and American Association of School Administrators, Educational Policies Commission. *Education and the Disadvantaged American.* Washington, D.C.: the Commission, 1962. 39 pp.

National Society for the Study of Education. *Social Forces Influencing American Education.* Sixtieth Yearbook, Part II. Chicago: University of Chicago Press, 1961. 252 pp.

Newman, Robert Ernest. *History of a Civic Education Project Implementing the Social-Problems Technique of Instruction.* Doctor's thesis. Palo Alto, Calif.: Stanford University, 1961. 492 pp. Abstract: *Dissertation Abstracts* 21: 3347; No. 11, 1961.

Palmer, John R. "History Textbooks and Social Change." *Social Education* 25: 135-36; March 1961.

RIESSMAN, FRANK. *The Culturally Deprived Child.* New York: Harper & Brothers, 1962. 140 pp.

ROBINS, L. N.; GYMAN, HARRY; and O'NEAL, P. "The Interaction of Social Class and Deviant Behavior." *American Sociological Review* 27: 480-92; August 1962.

ROSSI, PETER H., and DENTLER, ROBERT A. *The Politics of Urban Renewal: The Chicago Findings.* New York: Free Press of Glencoe, 1961. 308 pp.

SCHMIDT, L. B. "Function and Control of Pressure Groups." *Social Studies* 53: 55-59; February 1962.

SMITH, ERNEST A. *American Youth Culture: Group Life in Teenage Society.* New York: Free Press of Glencoe, 1962. 264 pp.

TODD, LEWIS PAUL, editor. "Peace Through World Law." *Social Education* 26: 361-434; November 1962.

TRAVERS, JOHN F., and DAVIS, RUSSELL G. "A Study of Religious Motivation and Delinquency." *Journal of Educational Sociology* 34: 205-20; January 1961.

TUMIN, MELVIN M. *An Inventory and Appraisal of Research on American Anti-Semitism.* New York: Anti-Defamation League of B'nai B'rith, 1961. 185 pp.

TUMIN, MELVIN M., editor. *Segregation and Desegregation: A Digest of Recent Research.* New York: Anti-Defamation League of B'nai B'rith, 1957. 112 pp.

U.S. COMMISSION ON CIVIL RIGHTS. *Education: 1961 Commission on Civil Rights Report.* Washington, D.C.: Government Printing Office, 1961. Book 2, 254 pp.

U.S. COMMISSION ON CIVIL RIGHTS. *Civil Rights U.S.A., Public Schools, Cities in North and West, 1962.* Washington, D.C.: Government Printing Office, 1962. 309 pp. (a)

U.S. COMMISSION ON CIVIL RIGHTS. *Civil Rights U.S.A., Public Schools, Southern States, 1962.* Washington, D.C.: Government Printing Office, 1962. 217 pp. (b)

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE, OFFICE OF EDUCATION. *The Impact of Urbanization on Education.* Summary report of a conference, May 28-29, 1962. Washington, D.C.: Government Printing Office, 1962. 12 pp.

WESTOFF, CHARLES F., and OTHERS. *Family Growth in Metropolitan America.* Princeton, N.J.: Princeton University Press, 1961. 433 pp.

CHAPTER IV

Psychological Forces Influencing Curriculum Decisions

NORMAN D. BOWERS

WHAT ARE the psychological forces that influence curriculum decisions? The immediate and traditional response of many educators includes three areas related to the pupil: principles of learning, characteristics of physical growth and development, and personality factors. The inference customarily made is that possession of accurate information in each area would aid one in making valid decisions about the curriculum.

However, other answers to this question are possible; and an alternative, although seemingly more complicated, was selected. In addition to the characteristics of learners and the learning process, psychological factors that influence curriculum decisions were considered to include the situational variables that are manipulated in the school environment to promote learning as well as the skills and the abilities that accrue to the learner as a result of education.

This chapter will be concerned with recent research that gives information about four areas: the characteristics of learners, the process of learning, the effects of different tasks and methods, and the definition of various skills and abilities that can be viewed as outcomes of school learning. It is intended that this chapter not only will direct attention to the already existing and extensive reviews of psychological research but also will evaluate some current research that appears promising. Moreover, it will suggest areas in which more research or a different approach appears appropriate. The distinction between social and psychological forces is arbitrary at times. Chapter III of this REVIEW is concerned with many of the forces that could be classified as psychological within the framework of an alternative organization.

The Learner

Extensive and numerous studies about various learner behavior characteristics exist. Among these, a considerable synthesis of information was apparent in the Sixty-Second Yearbook of the National Society for the Study of Education (1963), which dealt with child psychology. The lengthy bibliographies, which were available in the *Annual Review of Psychology*, edited by Farnsworth (1961-63), represented a valuable and useful source of research studies in developmental psychology. Specialized reviews concerned with the assessment of anxiety (Martin, 1961) and consideration of the self-concept (Lowe, 1961) offered not only a collection of research in these areas but also some evaluative comment. The February 1963 issue of the REVIEW summarized research related to the exceptional child. The availability of these publications should be helpful in the derivation of accurate summary statements that relate to learner characteristics.

In addition to these reviews, three different kinds of publications were of special interest relative to curriculum decisions. The first of these, *Growing Up in River City* (Havighurst and others, 1962), represented a study of certain features of the maturation process in a group of 487 individuals over a period of nine years. Concerned primarily with social institutions and their impact on adolescent behavior, the psychological factors that were studied included talent, personal and social adjustment, and school progress and achievement. Although relatively simple analyses of the data were given, a final report undoubtedly will offer additional insights that can be derived from this study.

The 1962 Yearbook of the Association for Supervision and Curriculum Development (1962) purported to describe how schools may help to develop sufficient, adequately functioning, and self-actualized persons. Containing much of interest and many statements on which there could be almost universal agreement, material derived initially from prepared statements of well-known authorities was cast within the framework of phenomenology. Although many interesting insights were given, little new quantitative evidence was presented.

The third, somewhat longer publication was the Sixty-First Yearbook of the National Society for the Study of Education (1962), which dealt with the perennial topic of individual differences. Various topics were summarized and discussed, and helpful bibliographies of research were given. This important area of individualizing instruction certainly remains one of the primary psychological forces that affect curriculum decisions.

The Learning Process

Of the numerous learning investigations completed during any given time, few seem to have direct relevance to the school curriculum area. After a review of the studies related to learning that were reported in the *Journal of Experimental Psychology*, it is hoped that more symposiums, such as the one edited by Carton and Hively (1959) in which Spence (1959), Melton (1959), and Underwood (1959) participated, will be held. In this symposium, application of laws, or principles, of learning to the classroom and to the school situation were considered. In the papers presented, it was proposed that a mutually supporting relationship be built between the science of learning and the technology of educational method. An illustration was given of the relevancy of a laboratory psychological variable to a school situation.

Two recent longer books are noteworthy in summarizing material and presenting different insights about human learning, learning theory, and teaching. Edited by Glaser (1962), the first book, which distinguished between training and education, reported the results of numerous military research programs involving the learning of complex skills. Although the point of view was somewhat different from that usually encountered in many curriculum publications, such as those prepared by the Association for Supervision and Curriculum Development (1959, 1960), curriculum workers not only should take heart, but also take heed of the suggestions made by interested and obviously competent researchers. The second publication offered a survey and interpretation of current psychological theory (Hill, 1963). Together with a clearly written contrast of learning in the school situation and of learning in the laboratory, a most informative essay on the present-day value of learning theory was included.

An interesting exchange of considered opinions that were related to verbal learning took place between Arnstine (1962) and Ausubel (1962). Such controversies, which are rarely published in the educational literature, should happen more frequently because clarification of one's ideas about the nature of the learning process is afforded.

Representative studies about learning with implications for curriculum decisions included those of Klausmeier and Check (1962), Lockard and Sidowski (1961), and Bilodeau, Sulzer, and Levy (1962). Continued study of the learning process in the laboratory and in applied situations is certain to continue. For educators, there exists a real need for the research to be forged into some kind of meaningful scheme. Clarification of curriculum theory might serve to simplify the task of selecting and analyzing important and relevant atomistic psychological studies related to learning. Work such as that of Beauchamp (1961) would provide a starting point for such an analysis.

The School Learning Situation

Although psychological factors sometimes are confined to the learner and to the process of learning, other important variables of a psychological nature exist within the situation of the school. The teacher and his constellation of traits represent one dimension of such variation. The basis for selecting among many alternatives that are related to tasks and methods suggests a second domain pertinent to psychological factors. The entire area of programed instruction and teaching machines raises questions about tasks and methods. Group relationships form another important variable within the situation. Obviously, the writer cannot do more than point out the general availability of research in these different areas.

The Teacher

The long-awaited *Handbook on Research in Teaching* (Gage, 1963) contains many references to the recent research literature concerned with the teacher and his role and activities in the classroom. This volume, which undoubtedly will become a standard reference for research workers interested in teaching and school learning, gives the most extensive and complete summary of research currently available on teachers and teaching.

Task and Method Variables

Since the publication of the December 1961 Review on growth, development, and learning, relatively little has appeared that offers conflicting or different results concerning task and method variables. Two investigations are typical of the kind of research in this area. Ausubel and Fitzgerald (1962) reported a study of the effects of advance organizers, antecedent learnings, and general background knowledge in learning and retention. Organizers enhanced learning for individuals with low verbal ability, and a general background knowledge also seemed to facilitate learning. In their study of performance after failure as a function of interpolated activity, Smoot and Lawson (1962) concluded, among other things, that rest after failure most effectively altered the influence of failure.

Two investigations seemed to emphasize the importance of introducing appropriate task and method variables in the preparation of teachers. Wittrock (1962) concluded that set operated on the behaviors of teacher candidates. Davis and Bowers (1961) reported that data that indicated training of future teachers in terms of specific kinds of learnings appeared to have positive effects. If task and method variables were inoperative or unimportant to either of these studies, it would appear that the results described could not have been demonstrated.

Automated Instruction and Teaching Machines

Numerous questions related to both psychological forces and curricular decisions have been raised by the recent upswing of interest in automated instruction, self-teaching devices, and programed materials. In the April 1962 Review, Silberman concluded that a greater burden of responsibility for learning would be placed on the quality of the instructional materials. Other recent reviews, including one by Morrill (1961), listed different areas necessary for a research program. Some of these problems have, of course, been dealt with in research studies. An entire issue of the *Journal of Educational Research*, edited by Lambert (1962), was devoted to describing a number of specific studies on one or more of these topics.

A publication from the military setting reported a conference dealing with programed learning jointly sponsored by the Office of Naval Research and the System Development Corporation (Coulson, 1962). This volume, which offered additional reviews of research as well as papers related to the theory of programed learning and to the factors important to computer-based instructional systems, gave a description of a computer-based laboratory for research and development in education. The summary of implications of current research and future trends by Stolurow (1962) constituted an interesting and hopeful picture of developments in this area. However, one also should read the paper by Pressey (1963), who pointed out that current programmers in blindly pursuing their art have been misled by a theory derived from comparative studies. He contended further that write-in responses were not as valuable as objective items; that reinforcement was not the significant process; and that programming was destructive of properties underlying a cognitive structure.

Group Interaction

Consideration of the school learning situation would be incomplete without some mention of the group forces operant within the school situation. A series of readings has been published by the National Training Laboratories (Bradford, 1961) as *Human Forces in Teaching and Learning.* This publication brought together a group of thoughtful and useful papers in one volume.

Important reviews of research studies on group interaction have been written by Mann (1959) and by Glanzer and Glaser (1959, 1961). The Fifty-Ninth Yearbook of the National Society for the Study of Education (1960) also constituted a helpful reference source of material upon group processes. Many individual studies continue to be reported in the periodic literature. It seems that the substantial information already available will make possible the effecting of some valid curriculum decisions.

Learning Outcomes

Given information about the pupil, the learning process, and the complexity of situational variables, the generalized outcome of education appears to be the development of the intellective process, including the so-called higher mental processes such as thinking, problem solving, and creative thinking that can be described in a theoretical framework such as that posed by Guilford (1959). Other neglected but important outcomes include such skill areas as reading, listening, and oral and written expression.

In their investigation of divergent thinking abilities and teacher ratings, Klausmeier, Harris, and Ethnathios (1962) found that the nature of the divergent thinking factors and the expression of them in classroom activities varied markedly according to sex and subject matter field. In an extensive study, Merrifield and others (1962) reported that no evidence of a unitary problem-solving ability was located—a finding which seemed to suggest that problem solving can be conceptualized in terms of factors in the structure of intellect. A noteworthy review of recent research on problem solving was supplied by Duncan (1959), who concluded, among other things, that the field of problem solving was poorly integrated. Perhaps, as integration is achieved, an increasing number of teachers and curriculum workers will become aware of the challenges and the issues posed by the concern that psychologists exhibit regarding the nature of human thinking and problem solving.

In reporting upon the results of three experiments dealing with the abstraction process, Ericksen (1962) suggested the importance of generalizing the abstraction process beyond the laboratory. It was contended that significant educational benefits would accrue from improved abstract learning abilities.

Much research on creative thinking has been completed, and a considerable portion has been included in two books by Torrance (1962) and by Getzels and Jackson (1962). Taylor and Holland, in the February 1962 Review, furnished a summary of studies relating to measurement of creativity.

In such outcome areas as reading, listening, and oral and written expression, some reviews of recent research exist, but relatively few individual studies are available. Harris (1963) summarized the investigations related to reading, and Strom (1963) surveyed investigations in the language arts area. Petty (1962) gave some research direction to the study of listening skill; Brown (1962) developed an index of listening; and Andrews (1962) described a comparative study of methods of developing listening ability. Other skill areas received little or no research attention in terms of well-designed, quantitatively oriented studies.

In an extensive effort, Kelley and others (1962) studied the development of cooperation in a minimal social situation. Although they discussed some possible implications of the research, the authors did not place their findings within the context of the school situation. Such an attempt to interpret their findings might be useful to educators.

It is obvious that, although there is a great deal of research available with respect to many lines of inquiry, many areas that have been neglected need attention and clarification. Little research related to learning outcomes has dealt with the comparative effect of task and method variables on learning. It is in this area, probably, that research workers can make lasting contributions.

Research Designs and Psychological Influences

It is rather obvious, after reading numerous individual research studies and reviews of research, that the task of integrating information is complex and difficult. At the same time, continued development in research-design efficiency seems to be dictated. Many studies showed a certain naïveté about research design and analysis. Fortunately, a number of recent publications dealing with this important topic should help to develop future sophistication. Campbell and Stanley (1963) wrote a lengthy and provocative chapter in the *Handbook on Research in Teaching* that was concerned with experimental and quasi-experimental designs. This chapter should prove to have a great effect on the quality of future research, as should the chapter by Medley and Mitzel (1963), who presented, among other things, a lucid treatment of analysis of variance involving complex components.

Other complex techniques for analysis of data seem to hold particular promise for work in the curriculum area. Chief among these are the multivariate methods that simultaneously consider a number of measurements on the same person. These procedures appear to hold the key to an accurate analysis of the range of abilities as well as of the skills or learning outcomes encountered in school situations. An introduction to multivariate techniques may be found in a book by Kendall (1957). Some explanation and helpful listings of FORTRAN-coded computer programs were given in a recent book on multivariate procedures by Cooley and Lohnes (1962).

A book by Winer (1962) containing many different designs should find wide usage among curriculum research workers. Particularly helpful was the detailed explanation of the meanings of interaction effects; the discussions related to the use of orthogonal components in tests for trend; and a series of illustrations and worked examples of nested factors, or hierarchal designs. This latter design, which is an often neglected analysis, is potentially useful because it can show the significance of differences for groups within methods, schools within cities, classes within schools, and the like.

The papers by Kempthorne (1961) and Stanley (1961), which appeared in the Phi Delta Kappa symposium on research, should be particularly helpful in planning and executing research involving psychological factors.

Conclusions

Although the research in the area of psychological factors appeared to be extensive and the reviews of the research were numerous, much of the available information did not seem to have direct usefulness and applicability in the making of curriculum decisions. Some direct attack on this problem is needed. In addition, at least two questions should be raised regarding the currently available data:

1. How does knowledge about psychological characteristics of learners become translated into patterns of action related to the curriculum?
2. In what ways would any given learning theories result in curriculum differences, including a differential in learning outcomes?

Although the newer statistical techniques may furnish valuable insights relative to the interpretation of complex data, development of a theory of curriculum is an important and needed concomitant to the realization of a coherent understanding of the implications to classroom learning that are present in the extensive psychological literature. The upswing in interest in educational problems, which has been evidenced by research workers in certain areas of psychology, should make the next few years productive in delineating psychological factors appropriate for curriculum decisions.

Bibliography

ANDREWS, DOROTHY E. "Comparative Study of Two Methods of Developing Music Listening Ability in Elementary School Children." *Journal of Research in Music Education* 10: 59-64; Spring 1962.

ARNSTINE, DONALD G. "Assumed Goals for Problem-Solving and Education: Chinks in the Defense of Verbal Learning." *Educational Theory* 12: 226-29; October 1962.

ASSOCIATION FOR SUPERVISION AND CURRICULUM DEVELOPMENT. *Learning More About Learning.* (Edited by Alexander Frazier.) Papers and reports from the Third ASCD Research Institute. Washington, D.C.: the Association, a department of the National Education Association, 1959. 88 pp.

ASSOCIATION FOR SUPERVISION AND CURRICULUM DEVELOPMENT. *Freeing Capacity To Learn.* (Edited by Alexander Frazier.) Papers and reports from the Fourth ASCD Research Institute. Washington, D.C.: the Association, a department of the National Education Association, 1960. 97 pp.

ASSOCIATION FOR SUPERVISION AND CURRICULUM DEVELOPMENT. *Perceiving, Behaving, Becoming: A New Focus for Education.* 1962 Yearbook. Washington, D.C.: the Association, a department of the National Education Association, 1962. 256 pp.

AUSUBEL, DAVID P. "In Defense of Verbal Learning: Beams in the Eyes of a Critic." *Educational Theory* 12: 230-33; October 1962.

AUSUBEL, DAVID P., and FITZGERALD, DONALD. "Organizer, General Background, and Antecedent Learning Variables in Sequential Verbal Learning." *Journal of Educational Psychology* 53: 243-49; December 1962.

BEAUCHAMP, GEORGE A. *Curriculum Theory.* Wilmette, Ill.: Kagg Press, 1961. 149 pp.

BILODEAU, EDWARD A.; SULZER, JEFFERSON L.; and LEVY, C. MICHAEL. "Theory and Data on the Interrelationships of Three Factors of Memory." *Psychological Monographs*: Vol. 76, No. 20 (Whole No. 539), 1962. 22 pp.

BRADFORD, LELAND P., editor. *Human Forces in Teaching and Learning.* Selected Reading Series Three. Washington, D.C.: National Training Laboratories, National Education Association, 1961. 102 pp.

BROWN, CHARLES T. "Introductory Study of Breathing as an Index of Listening." *Speech Monographs* 29: 79-83; June 1962.

CAMPBELL, DONALD T., and STANLEY, JULIAN C. "Experimental and Quasi-Experimental Designs for Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. Chapter 5, pp. 171-246.

CARTON, AARON S., and HIVELY, WELLS, II, editors. "A Symposium: 'Can the Laws of Learning Be Applied in the Classroom?'" *Harvard Educational Review* 29: 83-117; Spring 1959.

COOLEY, WILLIAM W., and LOHNES, P. R. *Multivariate Procedures for the Behavioral Sciences.* New York: John Wiley & Sons, 1962. 211 pp.

COULSON, JOHN E. *Programmed Learning and Computer-Based Instruction.* Proceedings of the Conference on Application of Digital Computers to Automated Instruction, October 1961. New York: John Wiley & Sons, 1962. 291 pp.

DAVIS, O. L., JR., and BOWERS, NORMAN D. "Group Sensitivity Training in a Teacher Education Program: An Initial Attempt." *Peabody Journal of Education* 39: 68-74; September 1961.

DUNCAN, CARL P. "Recent Research on Human Problem Solving." *Psychological Bulletin* 56: 397-429; November 1959.

ERICKSEN, STANFORD C. "Studies in the Abstraction Process." *Psychological Monographs*: Vol. 76, No. 18 (Whole No. 537), 1962. 28 pp.

FARNSWORTH, PAUL R., editor. *Annual Review of Psychology.* Palo Alto, Calif.: Annual Reviews. Vol. 12, 1961, 533 pp. Vol. 13, 1962, 602 pp. Vol. 14, 1963, 454 pp.

GAGE, N. L., editor. *Handbook of Research on Teaching.* A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. 1218 pp.

GETZELS, JACOB W., and JACKSON, PHILIP W. *Creativity and Intelligence.* New York: John Wiley & Sons, 1962. 293 pp.

GLANZER, MURRAY, and GLASER, ROBERT. "Techniques for the Study of Group Structure and Behavior: I. Analysis of Structure." *Psychological Bulletin* 56: 317-32; September 1959.

GLANZER, MURRAY, and GLASER, ROBERT. "Techniques for the Study of Group Structure and Behavior: II. Empirical Studies of the Effects of Structure in Small Groups." *Psychological Bulletin* 58: 1-27; January 1961.

GLASER, ROBERT, editor. *Training Research and Education.* Pittsburgh: University of Pittsburgh Press, 1962. 596 pp.

GUILFORD, J. P. "Three Faces of Intellect." *American Psychologist* 14: 469-79; August 1959.

HARRIS, THEODORE L. "Summary of Investigations Relating to Reading, July 1, 1961, to June 30, 1962." *Journal of Educational Research* 56: 283-301; February 1963.

HAVIGHURST, ROBERT J., and OTHERS. *Growing Up in River City.* New York: John Wiley & Sons, 1962. 189 pp.

HILL, WINIFRED F. *Learning: A Survey of Psychological Interpretations.* San Francisco: Chandler Publishing Co., 1963. 227 pp.

KELLEY, HAROLD H., and OTHERS. "The Development of Cooperation in the 'Minimal Social Situation.'" *Psychological Monographs*: Vol. 76, No. 19 (Whole No. 538), 1962. 19 pp.

KEMPTHORNE, OSCAR. "The Design and Analysis of Experiments with Some Reference to Educational Research." *Research Design and Analysis: Second Annual Phi Delta Kappa Symposium on Educational Research.* (Edited by Raymond O. Collier, Jr., and Stanley M. Elam.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1961. Chapter 4, pp. 97-133.

KENDALL, MAURICE G. *A Course in Multivariate Analysis.* New York: Hafner Publishing Co., 1957. 185 pp.

KLAUSMEIER, HERBERT J., and CHECK, JOHN. "Retention and Transfer in Children of Low, Average, and High Intelligence." *Journal of Educational Research* 55: 319-22; April 1962.

KLAUSMEIER, HERBERT J.; HARRIS, CHESTER W.; and ETHNATHIOS, ZACKARIA. "Relationships Between Divergent Thinking Abilities and Teacher Ratings of High School Students." *Journal of Educational Psychology* 53: 72-75; April 1962.

LAMBERT, PHILIP, issue editor. "The Teacher and the Machine." *Journal of Educational Research* 55: 405-531; June-July 1962.

LOCKARD, JOAN, and SIDOWSKI, JOSEPH B. "Learning in Fourth and Sixth Graders as a Function of Sensory Mode of Stimulus Presentation and Overt or Covert Practice." *Journal of Educational Psychology* 52: 262-65; October 1961.

LOWE, C. MARSHALL. "The Self-Concept: Fact or Artifact?" *Psychological Bulletin* 58: 325-36; July 1961.

MANN, RICHARD D. "A Review of the Relationships Between Personality and Performance in Small Groups." *Psychological Bulletin* 56: 241-70; July 1959.

MARTIN, BARCLAY. "The Assessment of Anxiety by Physiological Behavioral Measures." *Psychological Bulletin* 58: 234-55; May 1961.

MEDLEY, DONALD M., and MITZEL, HAROLD E. "Measuring Classroom Behavior by Systematic Observation." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. Chapter 6, pp. 247-328.

MELTON, ARTHUR W. "The Science of Learning and the Technology of Educational Methods." *Harvard Educational Review* 29: 96-109; Spring 1959.

MERRIFIELD, P. R., and OTHERS. "The Role of Intellectual Factors in Problem Solving." *Psychological Monographs:* Vol. 76, No. 10 (Whole No. 529), 1962. 21 pp.

MORRILL, CHARLES S. "Teaching Machines: A Review." *Psychological Bulletin* 58: 363-75; September 1961.

NATIONAL SOCIETY FOR THE STUDY OF EDUCATION. *The Dynamics of Instructional Groups.* Fifty-Ninth Yearbook, Part II. Chicago: University of Chicago Press, 1960. 286 pp.

NATIONAL SOCIETY FOR THE STUDY OF EDUCATION. *Individualizing Instruction.* Sixty-First Yearbook, Part I. Chicago: University of Chicago Press, 1962. 337 pp.

NATIONAL SOCIETY FOR THE STUDY OF EDUCATION. *Child Psychology.* Sixty-Second Yearbook, Part I. Chicago: University of Chicago Press, 1963. 550 pp.

PETTY, WALTER T. "Listening: Directions for Research." *Elementary English* 39: 574-77; October 1962.

PRESSEY, SIDNEY L. "Teaching Machine (and Learning Theory) Crisis." *Journal of Applied Psychology* 47: 1-6; February 1963.

SMOOT, KENNETH E., and LAWSON, REED. "Performance After Failure as a Function of Interpolated Activity." *Journal of Genetic Psychology* 100: 205-13; June 1962.

SPENCE, KENNETH W. "The Relation of Learning Theory to the Technology of Education." *Harvard Educational Review* 29: 84-95; Spring 1959.

STANLEY, JULIAN C. "Studying Status vs. Manipulating Variables." *Research Design and Analysis: Second Annual Phi Delta Kappa Symposium on Educational Research.* (Edited by Raymond O. Collier, Jr., and Stanley M. Elam.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1961. Chapter 6, pp. 173-208.

STOLUROW, LAWRENCE M. "Implications of Current Research and Future Trends." *Journal of Educational Research* 55: 519-27; June-July 1962.

STROM, INGRID M. "Summary of Investigations Relating to the English Language Arts in Secondary Education: 1961-1962." *English Journal* 52: 118-36; February 1963.

TORRANCE, E. PAUL. *Guiding Creative Talent.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 278 pp.

UNDERWOOD, BENTON J. "Verbal Learning in the Educative Processes." *Harvard Educational Review* 29: 107-17; Spring 1959.

WINER, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill Book Co., 1962. 672 pp.

WITTROCK, M. C. "Set Applied to Student Teaching." *Journal of Educational Psychology* 53: 175-80; August 1962.

CHAPTER V

Curriculum Components and Organization

RICHARD A. GIBBONEY, JOHN H. SANDBERG, DONALD M. CARROLL,
STANLEY N. MILLER, and NORMAN A. MILLER

THIS CHAPTER is concerned with the evaluative research relating to curriculum components and organization that was completed during the period covered by this issue of the REVIEW. Unimaginative descriptive reports and many studies uncritically designed and employing less than rigorous analytical techniques have been excluded. In preparing this chapter, the reviewers followed the procedures described below.

After the standard compilations of educational research were surveyed, a tentative bibliography was compiled. This bibliography was supplemented by telephone conversations with persons actively engaged in major curriculum revision projects as a check on the literature survey. Standards were established by which each study in the tentative bibliography was judged. These standards included such criteria as the relevance of the problem to current curriculum development, the adequacy of the experimental design, and the appropriateness of the statistical treatment of the data.

Each study in the bibliography was read in its entirety by the reviewers and by experimental researchers * on the staff of the Pennsylvania Department of Public Instruction, who applied the standards previously developed. Through this procedure, most of the studies in the tentative bibliography were eliminated. Those that remained are reviewed in this chapter.

Mathematics

There were few significant curriculum studies in mathematics. Those who seek more comprehensive reviews of recent research are referred to the surveys by Brown and Kinsella (1960), Brown and others (1963), Feierabend (1960), and Kellogg and Johnson in the June 1961 REVIEW. It should be noted that these writers were generally unenthusiastic about the research they reviewed. For instance, Brown and Kinsella expressed reservations concerning the narrowness of the scope of, the limited duration of, the small number of participating teachers in, and the restricted areas represented by most of the studies reported to them. Feierabend noted that many studies were also poorly designed.

The present chapter is confined to a few recent research efforts involving some of the School Mathematics Study Group (SMSG) courses.

---

* The assistance of Robert Hayes and Emanuel Berger is gratefully acknowledged.

While the studies reported here were not free from defects, they were significant attempts to contribute to knowledge in an important curriculum area. They also served to illustrate some of the difficulties inherent in conducting good curriculum research.

Minnesota National Laboratory Studies

The Minnesota National Laboratory began a series of SMSG studies in 1959. Reports have been issued periodically by the Minnesota State Department of Education, and these have been summarized by Rosenbloom (1961, 1962). The major purpose of these studies, which involved thousands of students in grades 6-12, was to determine whether students in SMSG courses do as well as students in conventional courses on standardized mathematics achievement tests. In addition, attention was given to such matters as the effects of (a) SMSG courses on students of various ability levels, (b) certain teacher characteristics on learning, and (c) various grouping arrangements in the teaching of SMSG materials. Unfortunately, space permits neither adequate description of these studies nor detailed discussion of the findings.

Although several of the data are enigmatic, inconclusive, and even contradictory, the following points emerge as tentative conclusions: (a) SMSG students do not appear to suffer on achievement tests, such as the *Sequential Tests of Educational Progress: Mathematics (STEP)*; (b) SMSG materials appear to be suitable for students with wide ranges of ability; (c) SMSG texts do not seem to place unusual demands on students' reading ability; (d) in general, teachers can do a satisfactory job with SMSG materials without special training; and (e) there is no evidence that SMSG materials are better adapted to one type of student group than to another.

Rosenbloom (1962) provided a critical assessment of the Minnesota National Laboratory studies. He cited, among other things, (a) the lack of adequate measures of achievement for both conventional and SMSG mathematics, (b) the difficulty of controlling teacher variability, and (c) the lack of knowledge about the possible effects of one course upon another, when the same teacher teaches both. Probably, few would disagree with Rosenbloom's contention that a properly designed experiment has not yet been effected. However, his revelation that SMSG and the Minnesota National Laboratory have worked out detailed plans for a five-year study is a source of encouragement.

Additional Reference: Kraft (1962).

Educational Testing Service Study

Another major effort to evaluate student achievement in SMSG courses was carried out by the Educational Testing Service (ETS) during the

1960-61 school year. As reported by Payette (1961), this study, which involved 300 classes at five grade levels (7, 9, 10, 11, 12) in 75 schools, was intended to provide evidence relating to the following questions: (a) Does the SMSG curriculum detract from student achievement with respect to traditional mathematical skills? (b) Does the SMSG curriculum result in a measurable extension of developed mathematical ability beyond that of conventional mathematics instruction? (c) How effectively is the SMSG curriculum communicated to students at various levels of scholastic ability? The following conclusions were derived from the data: (a) In general, students exposed to conventional mathematics had neither a pronounced nor a consistent advantage over students exposed to SMSG mathematics on achievement tests, such as *STEP* mathematics; (b) SMSG students scored significantly higher on tests designed specifically to measure achievement in SMSG mathematics than did those students pursuing a conventional curriculum; and (c) students of a wide range of ability levels appeared to profit from SMSG courses. However, Rosenbloom (1962) criticized the ETS study on many of the same grounds he used in his evaluation of the National Laboratory studies.

Other SMSG Research

Two University of Minnesota doctoral dissertations provided additional data regarding student achievement with SMSG materials. In a well-designed but small-scale study, Shuff (1962) compared the achievement on standardized tests (*STEP* and *Cooperative Mathematics Test for Grades 7, 8, and 9*) of groups of seventh and eighth grade students who had studied either SMSG or traditional courses. Unlike the National Laboratory and the ETS studies, he found statistically significant differences in favor of the students who had followed the traditional course.

In a companion study, Williams (1962) compared the performance of ninth and tenth grade students who had previously studied either SMSG or traditional courses. He concluded that traditional algebra texts appeared to be at least as effective as the SMSG materials in developing mathematical competence in grade 9, as measured by the *Cooperative Algebra Test: Elementary Algebra Through Quadratics*, and that traditional and SMSG materials seemed equally effective in preparing students in grades 8 and 9 for succeeding mathematics courses, whether SMSG or traditional. Neither Williams nor Shuff found evidence that SMSG materials at these grade levels are more suitable for high-ability students than for others.

The researchers whose work has been reported here would probably be the first to admit that the surface has barely been touched insofar as evaluation of the SMSG courses is concerned. The studies carried out so far actually are not evaluations of the SMSG courses. They are, for the most part, evaluations of student achievement on traditional standardized tests after using SMSG materials. It is necessary, as Rosenbloom (1962) states, that decisions be made regarding the objectives of mathematics education and that adequate instruments for measuring achievement of these objectives be developed. Only when this has been done can research determine which approaches to the teaching of mathematics are the most effective.

Science

This section will review four carefully selected studies which investigated pertinent problems related to national efforts to improve the science curriculum. The paucity of well-designed research concerned with significant curriculum problems revealed a glaring deficiency in research efforts over the past three years, even after allowance was made for the fact that some of the national curriculum projects have only recently been finished or are in their final stages of development.

Physics

Since its inception in 1956, the Physical Science Study Committee (PSSC) course in physics has opened new vistas for curriculum development and has provided a stimulus toward collaborative revision which is being felt in most subject fields today. More research must be completed before it can be known whether this course will achieve its objectives. No significant study was found that attempted to measure the effectiveness of the PSSC course in terms of its objectives.

A well-planned and meticulously executed descriptive study by Trowbridge (1961) is helpful in comparing the objectives of the PSSC course with those of the traditional physics course. A further analysis was made of the PSSC materials to determine their relevancy to the stated course objectives.

The investigator cited 72 objectives for high school physics based on a thorough review of all the PSSC materials and interviews with the course author. A review of traditional physics textbooks and supplementary materials and a survey of the literature of the past 10 years provided source materials for the traditional objectives. Of the 72 objectives thus secured, 17 were unique to the PSSC course, 17 to the traditional course, and 38 were joint objectives. The reliability coefficient was satisfactory for the procedure followed in determining the objectives.

The unique PSSC objectives were related to such factors as the unified nature of the course, the emphasis on basic science rather than on applied science, the reduction in the number of topics covered, and the emphasis on the laboratory as the focal point of learning. The unique traditional objectives emphasized consumer knowledge, the scientific method, the application of physics to technology, and laboratory work for the purpose of verifying principles. Joint objectives were related to such areas as appreciation and interest, teaching techniques, and concept development.

A questionnaire was sent to random samples of 100 PSSC teachers and 100 traditional teachers to determine the degree to which these teachers subscribed to the 72 objectives in belief and practice. The return was 85 percent and 76 percent of the recipients of the questionnaire in the respective groups. (Fourteen returns were analyzed separately to avoid contamination of the data by teachers who had attended PSSC institutes but were teaching traditional courses.) Six selected conclusions of this study are as follows:

1. PSSC teachers did show significantly higher agreement with belief in the unique PSSC objectives than did traditional teachers in 64.5 percent of the cases, but significantly higher practice in only 47 percent of the cases. PSSC teachers did not strongly reject the unique traditional objectives when teaching practice was considered.
2. Traditional physics teachers showed a tendency to subscribe to the unique PSSC and joint objectives in addition to the unique traditional objectives.
3. PSSC materials adhered to the PSSC objectives and to the joint objectives on such factors as (a) presentation of subject matter and (b) learning and teaching methods, on the basis of the percentage of these objectives that had rating values above the mean rating value for all objectives.
4. Significant correlations indicated that both PSSC and traditional teachers strongly adhered to the unique objectives of their courses with respect both to agreement and to practice.
5. PSSC teachers and traditional teachers differed significantly with respect both to agreement and to practice in their responses to the 72 objectives.
6. Unique traditional objectives received little attention in the PSSC materials (mean rating value of 1.20 compared with the over-all mean of 2.62).

The minor conclusions of this study, which were concerned with the responses of the PSSC and of the traditional teachers with respect to size of school, teaching experience, and class size, are of doubtful value, because the data merely describe possible relationships. Moreover, the study was not designed to control relevant variables.

Chemistry

During the 1961-62 school year, Heath (1962a, b) completed two studies that compared the achievement of students taking newly developed chemistry courses with that of students taking traditional chemistry courses. The new courses were those developed by the Chemical Education Material Study (CHEM) and by the Chemical Bond Approach Project (CBAP).

The procedure and design of the two studies were similar. Control teachers were randomly selected from lists of secondary schools. The experimental and control teachers were categorized in a three-variable stratification table: geographic region, size of community, and sex of teacher. The experimental and control groups were comparable on most relevant demographic characteristics. Data were analyzed by the analysis of covariance.

Significant differences on the *Cooperative Chemistry Test* were found between the CHEM classes (adjusted mean, 31.9) and the traditional classes (adjusted mean, 37.7), with intelligence, as measured by the *Cooperative School and College Ability Tests* (*SCAT*), controlled. This finding is not necessarily undesirable because of the difference in content of the CHEM course.

The comparison of the differences between the two groups on the CHEM final examination mean scores, with the effects of *SCAT* removed, cannot be stated with precision because one of the basic assumptions for covariance was not met. The adjusted mean for the CHEM group was 25.7; for the traditional group, 13.2. However, the investigator stated that the difference between the adjusted means was "quite meaningful." He noted, from an analysis of the regression lines, that the CHEM group's advantage over the traditional group became increasingly larger in the upper scholastic range. The investigator concluded finally that the CHEM group demonstrated considerable skill on both the traditional and the CHEM final examinations.

The results of the Chemical Bond Approach Project (CBAP) study were similar to those of the CHEM study. The traditional chemistry students achieved a significantly higher mean level of performance on the *Cooperative Chemistry Test* than did the CBAP classes when the effects of *SCAT* were controlled. The respective adjusted means were 38.1 and 34.2. The comparison of the two groups on the CBAP final examination was in the same direction as that reported above for the CHEM group on the CHEM final examination. In this case, the adjusted mean for the CBAP group was 25.7; for the traditional group, 15.6.

Heath's candor in pointing out the limited value of his studies to other problems of curriculum development is commendable. Fewer studies are needed that deal with global "how much" questions, and more refined studies are needed that deal with the qualitative "in what way" questions. Such a development would be most refreshing as a leaven for the many mediocre doctoral dissertations read for possible inclusion in this chapter.

DeRose (1962) was concerned with the possible transfer, between CBAP chemistry and PSSC physics, of attributes associated with the process dimension of the two courses: Are knowledge, attitudes, and methods acquired in one course transferable to the other, as compared with the responses of students who were taking one of the new courses for the first time?

Because of the relative recency of the curriculums and the particular hypotheses tested in this study, the sample selection procedures were too involved to be adequately summarized; however, the reviewers accepted the procedures used as appropriate to the conditions in existence at the time the study was done. The CBAP group consisted of 179 students; the PSSC group, 91. The sample was drawn from nine schools (six public, two university-related, and one private institution) in nine states. The students were similar on such factors as age, sex, and previous mathematical and scientific experience. Intelligence was contolled by the covariance technique. Approximately one-half of the students in the PSSC and about one-half of the students in the CBAP groups had completed one year of traditional chemistry and one year of traditional physics, respectively; the other half of the PSSC group had completed CBAP chemistry, and the other half of the CBAP group had completed PSSC physics. Appropriate tests developed by the ETS were used as the criterion measures.

No significant differences were found in PSSC physics between students who had studied CBAP chemistry and students who had studied traditional chemistry. No significant differences were found in CBAP chemistry between students who had studied PSSC physics and students who had studied traditional physics.

Four factors should be considered in future studies that deal with problems similar to those investigated here: (a) realization of adequate sampling; (b) validation of the criterion measures; (c) establishment of validity of comparisons on the process dimension, as judged on such factors as objectives and course materials; and (d) control of teacher variables to insure that teaching is in terms of the objectives of the particular course.

English

There has been much ferment in the field of English language instruction. Numerous articles discussing the confusion over desirable content and effective teaching techniques have appeared in professional journals. Significant research studies, however, have been sparse. This is, undoubtedly, why Project English was instituted by the U.S. Office of Education. Steinberg's (1962) report of the first Project English conference in May 1962 gave some idea of the confusion that exists. The conference participants recommended that the following questions be studied: What grammar should be taught? What effect, if any, does the teaching of grammar have on the student's ability to read, write, and speak? How can teachers best motivate students of English?

Theory of Language Instruction

Russell (1962) also discussed the confusion and pointed out the need for a research point of view as opposed to the opinion-oriented and careless methods used by many critics and by several investigators of English teaching. Rinker (1962) agreed that confusion exists and listed several priorities in the study of English teaching. Rinker's priorities as well as the questions listed by Steinberg, however, serve to illustrate Russell's thesis that much of what is said about English teaching is based on opinion rather than on new research evidence.

The confusion is compounded by arguments between traditional grammarians and structural linguists. Some research studies failed to show a significant relationship between knowledge of traditional grammar and ability in reading, speaking, and writing. Structuralists have seized on this lack of research support for traditional grammar to argue that the linguistic approach, if substituted for traditional grammar in schools, might improve pupil performance in reading, speaking, and writing.

Traditionalists and structuralists do not agree on the same major issues within their respective groups. Thomas (1962) reported that traditional English grammar textbooks do not agree on the number and definitions of parts of speech. Thomas advocated a unification and simplification that may allow the traditionalists and the structuralists to join forces.

Review of Studies

In a thorough review of the pertinent literature, Alwin (1958) attempted to integrate the structural principles that have been identified by the major structural linguists. Her research revealed four devices or principles of structural grammar: (a) word order, (b) word forms, (c) function words, and (d) intonation. Using procedures similar to Alwin's, O'Donnell (1961) attempted to identify "essential structural principles." O'Donnell and Alwin's agreement on the major principles lends authority to their work. But the very fact that Alwin spent approximately 100 pages and O'Donnell 40 pages to develop and defend these essential structural principles gives an indication of the embryonic nature of the principles. In more mature fields, such principles would require no extensive justification.

Careful research into the relative merits of structural linguistic *versus* traditional grammar approaches has barely begun. O'Donnell (1961) also analyzed the relationship between the reading comprehension of pupils and their awareness of the structural patterns of English. No standardized tests designed to measure the understanding of the patterns of English structure were available. Much of O'Donnell's effort, therefore, was directed to the development of a 50-item test to measure recognition of structural relationships in English. Test reliability was established by standard statistical procedures, including four administrations of the test to a total of 302 college and high school students. Each administration of the test was followed by a statistical analysis to determine the level of difficulty and the discriminative value of each item. The test items were accordingly revised. The final form of the test was administered to 106 subjects in the senior class of Bowling Green High School, Bowling Green, Kentucky. The results yielded a Spearman-Brown split-half reliability coefficient of .88 and a Kuder-Richardson interitem consistency coefficient of .86. The lack of randomness in selection of subjects limits the applicability of the study.

This group of 106 subjects also took Test C1: Reading Comprehension, the *Iowa Grammar Information Test*, and the *Otis Quick-Scoring Mental Ability Tests*. Various Pearson product-moment correlations were computed from these data. The following major findings resulted: (a) a correlation coefficient of .75 between scores on the grammar test and scores on the structure test, (b) almost identical coefficients of .46 and .44 between scores on the grammar and structure tests and scores on the reading comprehension test, (c) a coefficient of .90 between scores on the grammar test and vocabulary scores, and (d) a coefficient of .46 between scores on the structure test and vocabulary scores. The correlation coefficients were significant beyond the .05 level.

This relatively well-designed and well-conducted study indicated that awareness of structural elements of English and knowledge of the elements of traditional grammar had a significant correlation with reading comprehension. The high correlation of knowledge of grammar with vocabulary deserves further investigation.

No instruction in linguistic analysis of English occurred in this study. The test attempted to measure awareness of structural patterns with subjects who had received traditional instruction in English grammar. Still to be answered are such questions as these: What would be the effect of varying amounts of instruction in linguistic analysis of English on reading comprehension and speaking and writing patterns? What is the relationship between writing and speaking patterns and awareness of structural English relationships?

It appears that some of those teachers and research workers concerned with English language instruction must think through the assumptions on which English language instruction is based. If the major objectives of language instruction are related to appreciating the complexity of language, to understanding the structural relationships of language, or to knowing its uses, such a curriculum may then be designed and evaluated. But, if instruction in English language is to lead to improved reading, speaking, and writing ability, much careful experimentation must be undertaken. Such investigations will examine alternative curriculum content along with alternative instructional methods and techniques in an effort to determine whether these abilities can, in fact, be improved significantly by instruction.

On the basis of the research evidence and journal articles reviewed, there is little reason to assume that current objectives for English language instruction are sound or attainable. Moreover, it cannot be assumed either that current curriculum content adds to these objectives or that current instructional techniques are efficient.

Foreign Languages

Most of the acceptable research in the teaching of foreign languages concerned the elementary school. The program of teaching foreign languages in the elementary school is referred to as FLES. A study by the Somerville Public Schools (1962) examined the effect of modern language study in the elementary school on achievement in high school. Somerville High School received pupils from elementary schools which offered a FLES program in grades 3-8 and from elementary schools which did not. After comparing the over-all high school grade point average of 1,530 FLES and non-FLES pupils, the investigators concluded that the FLES program did not have an adverse effect on the achievement of pupils of equal general ability.

A comparison of 343 pupils who elected the third level of foreign languages in high school revealed that the grade point average in languages studied was not significantly different when FLES and non-FLES pupils of the same general ability were compared. The FLES pupils were, however, accelerated one year in French and in Spanish.

Lopato's (1961) study, which was conducted at the third grade level with two FLES and two non-FLES classes, attempted to determine whether a FLES program had an adverse effect upon pupil achievement in other basic subjects. The experimental and control classes were not significantly different in intelligence or social class background.

Standardized achievement tests in the basic skill subjects were administered to both groups at the beginning and end of the experimental period. No significant difference in mean gain in achievement was found in favor of the control groups, although a significant difference in arithmetic achievement was found in favor of the experimental groups. Caution should be exercised in interpreting these findings, in view of the possible presence of both teacher variability and the Hawthorne effect.

The investigators concluded that children in a regular third grade class can furnish evidence of satisfactory achievement in conversational French when they have been taught by qualified classroom teachers. At the end of the year, each pupil was rated on comprehension, fluency, and pronunciation by a panel of judges. Reliability of the scale for rating French achievement was considered established if there was at least 67 percent agreement among the judges on each rating.

It was also concluded that pronunciation and fluency of the third grade children were superior to the results generally obtained with beginning high school students after an equivalent number of hours of instruction, although wide variations were still evident.

Social Studies

Several major developmental studies, which hopefully may give direction to the development of a modern social studies curriculum, are in their planning stages.

One study should be noted because it bears upon an important question that has been raised relative to the teaching of chronology in the elementary grades. Arnsdorf (1959) conducted an experiment with 563 sixth grade subjects to determine whether children of this level might be helped in their understanding of several measures of chronology through an instructional program which stresses activities dealing with chronology.

The investigator concluded that children at the sixth grade level might be helped in their understanding of several measures of chronology through an instructional program. A further conclusion was that the upper ability group appeared to be most successful in its performance on measures of chronology.

Additional References: Fraser (1962); *School Review* (1962).

Curriculum Organization

During the late 1950's and early 1960's, certain curriculum organizations were suggested which were basically different from those used in previous decades. Trump (1960), among others, emphasized such techniques as team teaching, large- and small-group instruction, and flexible scheduling. Goodlad and Anderson (1959) related continuous pupil progress to elementary school organization. These organizations were discussed with increasing frequency in the literature during the period from 1959 to 1963. Most of the literature, however, was not based on systematic research studies which evaluated the ideas presented. Two well-designed studies have been selected for review. Unfortunately, both apply to the secondary school.

Team Teaching

Taffel (1961) evaluated a team method of teaching a first course in high school physics to academically talented students, through comparing the achievement of the experimental classes taught by the team method with that of the control classes taught by the conventional single teacher method. The study also analyzed the attitudes of students and teachers toward the team method.

Fifty-five pairs of students were obtained by matching age, sex, grade, science-mathematics scores, intelligence, and scores on the *Dunning Physics Test*. One student from each pair was placed in the experimental class, and one student from each pair was placed in a corresponding control class. There were no significant differences between the two groups on the criteria used in matching.

A teaching team was formed consisting of the investigator, who was an experienced physics teacher and science supervisor, and two science teachers, who were comparable to each other in experience and education.

The investigator taught the large lecture group for the team twice each week. The team teachers taught the remainder of the team program in the conventional classes. Each of the team teachers also taught one of the control classes following the traditional pattern of the school. Time, content, rate of content coverage, and scheduling were the same for both groups except for the large group lectures received by the experimental classes.

The criterion measures were the *Dunning Physics Test*, the Bronx High School of Science midyear examination in physics, and the New York State Regents examination in physics. The investigator reported no significant differences in achievement between the experimental and control groups. The results of a questionnaire, however, indicated favorable attitudes by both students and teachers toward the team teaching approach. Taffel concluded that team teaching is a practical alternative to the conventional method of teaching physics to academically talented students.

The present reviewers believe that team teaching is characterized by cooperative planning and by small-group instruction in addition to other major elements. In this study, the time devoted to planning was minimal, and no small-group instruction was reported. A question may be raised, therefore, concerning whether the study was, in fact, an evaluation of team teaching, or, what appears more likely, an evaluation of large-group instruction.

Track Organization

McCown (1960) evaluated the influence of grouping practices on the achievement of students in the honors and in basic curriculum tracks of all senior high schools in the District of Columbia. That portion of the study concerning the honors track is reviewed here.

The District of Columbia initiated a four track curriculum in the tenth grade in September of 1956. McCown compared 275 honors track graduates of 1959 with a matched group of graduates from the class of 1958 who were heterogeneously grouped. Groups were matched on intelligence, reading level, and arithmetic achievement. In the matching procedure, equal numbers of honors track students and of heterogeneously grouped students were taken from six high schools to form the experimental and the control groups.

The *College Entrance Examination Board Scholastic Aptitude Test* and the *Iowa Tests of Educational Development* in English, social studies, natural science, and mathematics were given to the members of both experimental and control groups during their senior year. The differences on all measures were significant in favor of the honors track group. Other data were presented pertaining to Merit Scholarship semifinalists, class grades, and student and teacher opinion. The findings based on these data were not cited, because they were not related to the stated problems investigated, in the instance of student and teacher opinion, or were not treated statistically, in the case of class grades. Findings based on the Merit semifinalists are not reported, because the investigator included all scholarship holders in 1959, whether or not they had taken the honors track program; therefore, the increase in Merit semifinalists could not necessarily be attributed to the influence of the honors track program.

Recommendations

One conclusion must be stated briefly based on an analysis of studies reviewed for this chapter: Many doctorates are being awarded for poorly designed research about insignificant problems. That so few doctoral studies, of those read by the authors, were included in this chapter is eloquent testimony to this statement. This shocking lack of discrimination on the part of some faculties not only misleads the degree candidate but also retards the development of the profession by dulling its intellectual cutting edge.

Since few studies, doctoral or otherwise, are related to a theoretical framework of learning or teaching or to a comprehensive curriculum design, the separate studies, when viewed as a whole, appear as sounding brass and tinkling cymbals. Too few studies are replicated; too few studies have an end beyond themselves. One recommendation might be for more graduate schools to plan a series of doctoral studies dealing with one aspect of a significant problem. The general lack of a true residence requirement for doctoral students probably is related to the poor quality of the research surveyed.

This journal could perform a vital service by rigorously selecting fewer studies for review, by analyzing each study in depth, and by candidly reporting its strengths and weaknesses. Exclusion from the REVIEW would increasingly become the vehicle for vigorous internal professional criticism.

Bibliography

ALWIN, VIRGINIA VIVIAN. *A Study of the Possible Contributions of Modern Linguistic Science to the Teaching of the English Language in High School.* Doctor's thesis. Minneapolis: University of Minnesota, 1958. 211 pp. Abstract: *Dissertation Abstracts* 21: 1133; No. 5, 1960.

ARNSDORF, VAL EARL. *An Investigation of Teaching of Chronology in the Sixth Grade.* Doctor's thesis. Minneapolis: University of Minnesota, 1959. 162 pp. Abstract: *Dissertation Abstracts* 20: 3992-93; No. 10, 1960.

BROWN, KENNETH E., and KINSELLA, JOHN J. *Analysis of Research in the Teaching of Mathematics: 1957 and 1958.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1960, No. 8. Washington, D.C.: Government Printing Office, 1960. 50 pp.

BROWN, KENNETH E., and OTHERS. *Analysis of Research in the Teaching of Mathematics: 1959 and 1960.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1963, No. 12. Washington, D.C.: Government Printing Office, 1963. (In press)

DeRose, James V. *Relationships Between Achievement in PSSC Physics and CBAP Chemistry.* Doctor's thesis. Philadelphia: University of Pennsylvania, 1962. 163 pp.

Feierabend, Rosaline L. "Review of Research on Psychological Problems in Mathematics Education." *Research Problems in Mathematics Education.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Program, Monograph No. 3. Washington, D.C.: Government Printing Office, 1960. pp. 3-46.

Goodlad, John I., and Anderson, Robert H. *The Nongraded Elementary School.* New York: Harcourt, Brace and World, 1959. 248 pp.

Heath, Robert W. *A Study of Achievement in High School Chemistry: Report to the Chemical Bond Approach Project.* Princeton, N.J.: Educational Testing Service, 1962. 12 pp. (a)

Heath, Robert W. *A Study of Achievement in High School Chemistry: Report to the Chemical Education Material Study.* Princeton, N.J.: Educational Testing Service, 1962. 12 pp. (b)

Lopato, Esther Wolf. *An Experiment To Determine the Effect of Learning Conversational French on Academic Achievement of Third Grade Children.* Doctor's thesis. New York: New York University, 1961. 155 pp. Abstract: *Dissertation Abstracts* 22: 160-61; No. 1, 1961.

McCown, George Wilbur. *A Critical Evaluation of the Four Track Curriculum Program of the District of Columbia Senior High Schools with Recommendations for Improvement.* Doctor's thesis. College Park: University of Maryland, 1960. 177 pp. Abstract: *Dissertation Abstracts* 21: 2558-59; No. 9, 1961.

O'Donnell, Roy Cecil. *The Relationship Between Awareness of Structural Relationships in English and Ability in Reading Comprehension.* Doctor's thesis. Nashville: George Peabody College for Teachers, 1961. 119 pp. Abstract: *Dissertation Abstracts* 22: 1989-90; No. 6, 1961.

Payette, Roland F. "Educational Testing Service Summary Report of the School Mathematics Study Group Curriculum Evaluation." *Reports on Student Achievement in SMSG Courses.* Newsletter No. 10. Stanford, Calif.: School Mathematics Study Group, Stanford University, 1961. pp. 5-11.

Rinker, Floyd. "Priorities in the English Curriculum." *English Journal* 51: 309-12; May 1962.

Rosenbloom, Paul C. "Minnesota National Laboratory Evaluation of SMSG, Grades 7-12." *Reports on Student Achievement in SMSG Courses.* Newsletter No. 10. Stanford, Calif.: School Mathematics Study Group, Stanford University, 1961. pp. 12-26.

Rosenbloom, Paul C. *Effectiveness of the SMSG Material for Grades 7-12.* Publication No. XXXVIII-B-228(1). St. Paul: Minnesota State Department of Education, 1962. 7 pp. (Mimeo.)

Russell, David H. "Councilletter: Research in English—Dynamic or Deluding?" *English Journal* 51: 216-18; March 1962.

Shuff, Robert Vance. *A Comparative Study of Achievement in Mathematics at the Seventh and Eighth Grade Levels Under Two Approaches, School Mathematics Study Group and Traditional.* Doctor's thesis. Minneapolis: University of Minnesota, 1962. 221 pp. Abstract: *Dissertation Abstracts* 23: 558-59; No. 2, 1962.

Somerville Public Schools. *Evaluation of the Effect of Foreign Language Study in the Elementary School upon Achievement in the High School.* Somerville, N.J.: Board of Education, 1962. 85 pp.

Steinberg, Erwin R. "Needed Research in the Teaching of Secondary School English." *English Journal* 51: 617-20; December 1962.

Taffel, Alexander. *An Evaluation of a Team Method of Teaching High School Physics to Academically Talented Students.* Doctor's thesis. New York: New York University, 1961. 134 pp. Abstract: *Dissertation Abstracts* 22: 4297-98; No. 12, 1962.

Thomas, Owen. "Generative Grammar: Toward Unification and Simplification." *English Journal* 51: 94-99, 113; February 1962.

Trowbridge, Leslie W. *A Comparison of the Objectives of Traditional High School Physics with the Objectives of the Physical Science Study Committee Course, and an Analysis of the Instructional Materials of the Physical Science Study Committee Course.* Doctor's thesis. Ann Arbor: University of Michigan, 1961. 273 pp. Abstract: *Dissertation Abstracts* 22: 812; No. 3, 1961.

Trump, J. Lloyd. *New Directions to Quality Education: The Secondary School Tomorrow.* Washington, D.C.: National Association of Secondary-School Principals, a department of the National Education Association, 1960. 14 pp.

WILLIAMS, EMMET DAVID. *Comparative Study of SMSG and Traditional Mathematics.* Doctor's thesis. Minneapolis: University of Minnesota, 1962. 193 pp. Abstract: *Dissertation Abstracts* 23: 560; No. 3, 1962.

Additional References

FRASER, DOROTHY M. *Current Curriculum Studies in Academic Subjects.* Washington, D.C.: Project on Instruction, National Education Association, 1962. 102 pp.

KRAFT, CHARLES H. *Evaluation of SMSG Grades 7-12.* Publication No. XXXVIII-B-235. St. Paul: Minnesota State Department of Education, 1962. 10 pp.

SCHOOL REVIEW. "Curriculum Development Projects Supported by the National Science Foundation." *School Review* 70: 1-147; Spring 1962.

CHAPTER VI

Teaching

ROLLAND CALLAWAY and GERALD T. GLEASON

IN THE June 1960 REVIEW, which was devoted to curriculum planning and development, Bellack and Huebner presented a rationale for the inclusion of a chapter on teaching. They indicated that, although teaching is closely related to curriculum, little progress had been made in the direction of development of a theory of teaching as a vehicle for the implementation of curriculum. It is encouraging to note that significant progress has been made in this direction during the past few years. A number of studies included in the section on analysis of the teaching act report the development of theoretical constructs and instruments for the description and analysis of teaching.

Other studies have been concerned with an examination of the role of personality as well as with problems in teacher recruitment, selection, and effectiveness. Closely related are the studies on the characteristics of teachers in the broader personal, social, intellectual, and cultural aspects.

The reader who is interested in a comprehensive analysis of this topic will be aided considerably by three recent major publications. The *Handbook of Research on Teaching* (Gage, 1963) presents a scholarly and critical analysis of the literature, including sections on theoretical orientation, methodology of research on teaching, major variables, teacher personality and characteristics, and methods of teaching various grade levels and subject areas. This volume is likely to be an invaluable source for many years. An entire issue of the *Journal of Experimental Education* (Barr, 1961) summarized much of the research completed within the past decade on prediction of teacher success. Highly significant was Ryans' (1960) comprehensive study on teacher characteristics and personality. Each of these sources, which represent comprehensive and authoritative summaries and syntheses of research on teaching, should enable researchers and curriculum workers to achieve greater understanding and to develop penetrating insights into the nature of teachers and teaching.

Analysis of the Teaching Act

Experimental studies designed to develop methods and instruments for studying the teaching act have made encouraging progress during the period covered by this REVIEW. Medley and Mitzel (1963) summarized the research and attempts at measuring classroom behavior based on systematic observation. Such studies have been plagued by either lack of, or multiplicity of, a concept of teaching; by problems of instrumentation, design, and replication; and by the changing public and professional image of the ideal teacher.

Flanders (1961) continued to develop his interaction analysis, an observational technique which classifies verbal behavior into categories that have been used to analyze the degree and effects of direct and indirect teacher influence as well as a range of teacher and student behavior. Through the use of an ingenious analysis matrix, Flanders was able to specify and quantify a number of significant acts of the teacher that conceivably influence achievement and adjustment of the student. Flanders also used the method of interaction analysis as an in-service tool for helping teachers analyze and change their classroom behavior.

As part of a study on assessment of quality of teaching, Hughes (1962) compiled and analyzed verbatim records of a range of teachers judged successful and less successful. She identified 31 specific functions under seven broad classifications: control, imposition, facilitation, development of content, personal response, positive affectivity, and negative affectivity. These categories and functions, which were used to describe and classify verbal behavior in the classroom, were analyzed in terms of a number of significant aspects of content and curriculum implications.

Medley and Mitzel (1958) continued development of an instrument called the *Observation Schedule and Record*. Originally designed to provide quantitative data on behavior of beginning teachers, the instrument has been revised and refined to increase observer accuracy, increase efficiency of use, and separate the process of observing from the process of scoring. Factor analysis of the data yielded largest loadings on the factors of emotional climate, verbal emphasis, and social organization.

Smith (1961) focused directly upon the logical structure of a teacher's verbal behavior. Through separating the teaching act from learning and through analyzing the recordings of verbal interaction in high school classes, he identified two categories of teacher verbal behavior—performative and expressive. Wright (1959) developed an instrument for analysis of the verbal interaction of mathematics teachers, and Brown (1960) used the instrument designed by Withall (1951) to study the behavior of third grade teachers.

The problems associated with the development of instruments for measurement of classroom behavior were summarized by Medley and Mitzel (1963). At first glance, they would appear to be insurmountable. The reliability and validity of observations as a means of assessing teaching effectiveness were studied by Groff (1961), who found that college supervisors agreed with the student teacher's self-image only slightly more than one-half of the time. Veldman, Peck, and McGuire (1961) have suggested that the individual construct systems of judges or raters of human behavior are significant variables in the research process.

However, the promising insights these and other researchers have provided into the nature of the teaching act represent real progress toward an improved understanding of the task of teaching and the eventual development of a theory of instruction.

The Teacher and Teaching Effectiveness

Personality

Much of the research on teaching has been based on the premise that teaching effectiveness is a function of the teacher's personality. These studies have been concerned either with the total personality pattern or with specific personality traits as related to the prediction of teaching success or effectiveness. It should be recognized that any study concerned with the personality of the teacher is generally based on a stated or implied concept or on an operational definition of teaching. In this context, Barr (1960) asked whether the techniques of teaching that grow out of learning theory are encompassed by the behavioral aspects of teaching personality, or whether the behaviors found in the techniques of teaching are something different. Jensen (1961) suggested that there are a limited number of nonoverlapping critical qualities and that each quality is essential. Thus, the good teacher would possess at least one superior quality and would not fall below average on any others considered necessary.

A major problem in research on teaching has been the development of instruments or scales that measure personality patterns or traits and that might be correlated with teaching effectiveness. For example, Sheldon (1959) indicated that "fakability" should be considered as well as reliability and validity in the development and use of teacher-selection inventories.

Among the most popular instruments that have been utilized are the *Minnesota Teacher Attitude Inventory* (*MTAI*) and the *Guilford-Zimmerman Temperament Survey* (*GZTS*). Popham and Trimble (1960) hypothesized that the *MTAI* discriminates between superior and inferior teachers and concluded that the *MTAI* can be used as an indication of teachers' general competencies as well as an index of the types of social atmosphere that teachers maintain in the classroom. In another study, Cook (1962) compared *MTAI* and *GZTS* trait correlations for a group of beginning teachers, student teachers, and in-service teachers. Gruber (1960) also utilized the *MTAI* and the *GZTS* in a study concerned with the personality traits and attitudes of undergraduate majors in, and teachers of, physical education. Gowan (1960) attempted to develop a screening scale related to the *GZTS* as a prognosis of teaching effectiveness.

In the use of another instrument, the Taylor Manifest Anxiety Scale, Dutton (1962) concluded that student teachers can be differentiated in relation to their ability to be influenced and to their attitudinal changes toward youth. Lang (1960) used the *Edwards Personal Preference Schedule* and the Lang Scale of Motives for Teaching in his study of teacher selection. Scandrette (1962) also employed the *Edwards Personal Preference Schedule* to determine whether women students preparing to teach on the elementary level have different need patterns from students preparing to teach on the secondary level. Garrison and Scott (1961) used the same scale in a similar study.

The *Minnesota Multiphasic Personality Inventory* (*MMPI*) has also been administered in a number of studies concerned with teaching effectiveness. Flanagan (1961) compared scores on the *MMPI* with teaching success indicated by supervisory ratings. Utilizing the *MMPI* and the group Rorschach scale, Cole (1961) found that they correlated significantly with success in teaching as rated by a classroom observer.

An example of the self-report approach to research on teaching was Hatfield's (1961) study, which attempted to determine the relationship between self-concept and successful performance in student teaching. Also in this category was Peck's (1960) investigation in which he obtained personality analyses from sentence completion data.

Another approach to research on teaching is concerned with the study of specific personality traits. For example, Dixon and Morse (1961) found that teachers who have high emphatic capacity, as indicated by pupils, are seen also as better teachers. Reed (1961) summarized four independent studies concerned with teacher warmth and competency. In another investigation concerned with a specific personality trait, Dugan (1961) studied the egocentric teacher and altruistic teacher. Powell and Ferraro (1960) examined sources of tension in married and single women teachers.

Characteristics

In general, the studies concerned with characteristics of teachers represent attempts to relate either a specific attribute or a variety of characteristics of teachers to teaching effectiveness as measured by observer ratings or pupil gain. Perhaps the traditional approach to the investigation of teacher characteristics has been the development of lengthy and comprehensive lists of specific qualities that are judged to be related to teaching effectiveness. The study of Cassel and Johns (1960), in which they examined the critical characteristics of effective teachers as assessed by school administrators, is an example. Tyler (1960) noted the almost endless number of characteristics as one of the problems in this approach.

The work of Ryans (1960), as described in *Characteristics of Teachers*, represents one of the most comprehensive studies in this area. Ryans (1961b) also reported on a study of relationships between an observer's assessment of classes of pupils and teachers and the classroom behavior of pupils and teachers. In another study, Ryans (1961a) related certain teacher characteristics, as indicated by scores on a self-report inventory, to pupil behavior in the classroom and to pupil-teacher relationships.

In another comprehensive study in this category, Heil and Washburne (1962) hypothesized that there was a relationship between the changes that take place in children under the influence of teachers and the types of teachers and children involved. They found that there are identifiable types of teachers and that different types of teachers have different effects on the children they teach.

One of the persistent questions related to this category concerned the authoritarian or democratic behavior of the teacher. After reviewing 49 experimental studies in which authoritarian leadership was compared with democratic leadership, Anderson (1959) concluded that the evidence did not demonstrate that authoritarian or democratic leadership was consistently associated with higher productivity. In relation to this question, Johnston (1959) studied personality differences between autocratic and democratic teachers and concluded that democratic teachers had more basic social-emotional security. In another study, Christensen (1960) was concerned with the relationships between permissiveness and warmth of teachers and affect-need and achievement of pupils. Medley and Mitzel (1959) also considered behavioral correlates of teacher effectiveness.

A number of studies have been reported that were concerned with students' assessments of teacher behavior and effectiveness. At the college level, Drayer (1961) investigated students' views of the qualifications of their teachers. Likewise, at the college level, Cosgrove (1959) developed a technique for evaluating the effectiveness of the instructor by the student.

Included in the studies of teacher characteristics are those that compare the characteristics of teachers with those of a general population. For example, Mitzel and Dubnick (1961) reported on a persisting problem—the scholastic ability of prospective teachers as compared with that of liberal arts students. They indicated that perhaps teachers are not representative of the general scholastic ability available.

Finally, in relation to teacher characteristics, are those studies pertaining to general background, choice of teaching as a career, selection of teaching area or level, and persistence in teaching. In this category, Rabinowitz and Crawford (1960) made a comprehensive investigation of the teaching careers of graduates of the municipal colleges of New York City concerning factors related to entering teaching and persisting in teaching as a career. In a related study, Kleinman (1960) examined the adjustment that beginning teachers make in their first year of teaching in relation to the factors which were of concern to them as they selected their first teaching situation. In relation to selection of teaching level, Fox (1961) studied factors that influence the choice of elementary or secondary teaching, and Jones and Gottfried (1962) attempted to identify the interrelationships between preferences for teaching various types of exceptional children.

Conclusion

Review of the research on teaching during the past three years indicates a number of promising developments growing out of the work of several researchers on the nature of the teaching act. These studies might be described as micromatic approaches, because they attempt to identify and classify specific observed behaviors and use such classifications in an attempt to develop a theoretical base for the amorphous act of teaching. Similarly, the studies reported on teaching effectiveness and characteristics of teachers typically utilize an approach which assumes that such characteristics, while obviously interrelated, can be identified and isolated, at least statistically. Identification and cultivation of those characteristics that tend to be related to successful teaching can only lead in the direction of more successful predictions of success in teaching as well as appropriate adaptation of preparation programs.

The tasks remaining are frightening in their complexity. However, through continued research designed to move us from empirical evidence toward postulation of theoretical constructs and toward the systematic formulation of theories of instruction from which deductions may be tested, the profession will be able to define its role and responsibility effectively. From this perspective, significant beginnings seem to be taking shape—beginnings which can be a source of hope to the research-oriented educator.

Bibliography

ANDERSON, RICHARD C. "Learning in Discussions: A Résumé of the Authoritarian-Democratic Studies." *Harvard Educational Review* 29: 201-15; Summer 1959.

BARR, A. S. "The Assessment of the Teacher's Personality." *School Review* 68: 400-408; Winter 1960.

BARR, A. S., issue editor. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations." *Journal of Experimental Education* 30: 5-156; September 1961.

BROWN, GEORGE I. "Which Pupil to Which Classroom Climate?" *Elementary School Journal* 60: 265-69; February 1960.

CASSEL, RUSSELL N., and JOHNS, W. LLOYD. "The Critical Characteristics of an Effective Teacher." *National Association of Secondary-School Principals Bulletin* 44: 119-24; November 1960.

CHRISTENSEN, CLIFFORD M. "Relationships Between Pupil Achievement, Pupil Affect-Need, Teacher Warmth, and Teacher Permissiveness." *Journal of Educational Psychology* 51: 169-74; June 1960.

COLE, DAVID L. "The Prediction of Teaching Performance." *Journal of Educational Research* 54: 345-48; May 1961.

COOK, DESMOND L. "A Note on the Relationships Between MTAI and GZTS Scores for Three Levels of Teacher Experience." *Journal of Educational Research* 55: 363-67; May 1962.

COSGROVE, DON J. "Diagnostic Rating of Teacher Performance." *Journal of Educational Psychology* 50: 200-204; October 1959.

Dixon, W. Robert, and Morse, William C. "The Prediction of Teaching Performance: Emphatic Potential." *Journal of Teacher Education* 12: 322-29; September 1961.

Drayer, Adam M. "Students' Views of the Qualifications of Their Teachers." *Journal of Teacher Education* 12: 338-41; September 1961.

Dugan, Ruth R. "Personality and the Effective Teacher." *Journal of Teacher Education* 12: 335-37; September 1961.

Dutton, Wilbur H. "Attitude Change of Elementary School Student Teachers and Anxiety." *Journal of Educational Research* 55: 380-82; May 1962.

Flanagan, Carroll Edward. "A Study of the Relationship of Scores on the MMPI to Success in Teaching as Indicated by Supervisory Ratings." *Journal of Experimental Education* 29: 329-54; June 1961.

Flanders, Ned A. "Analyzing Teacher Behavior as Part of the Teaching-Learning Process." *Educational Leadership* 19: 173-75, 178-80; December 1961.

Fox, Raymond B. "Factors Influencing the Career Choice of Prospective Teachers." *Journal of Teacher Education* 12: 427-32; December 1961.

Gage, N. L., editor. *Handbook of Research on Teaching.* A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. 1218 pp.

Garrison, Karl C., and Scott, Mary Hughie. "A Comparison of the Personal Needs of College Students Preparing To Teach in Different Teaching Areas." *Educational and Psychological Measurement* 21: 955-64; Winter 1961.

Gowan, J. C. "A Teaching Prognosis Scale for the Guilford-Zimmerman Temperament Survey." *Journal of Educational Research* 53: 345-48; May 1960.

Groff, Patrick. "The Personality Self-Images of Student Teachers." *Journal of Teacher Education* 12: 433-36; December 1961.

Gruber, Joseph John. "Personality Traits and Teaching Attitudes." *Research Quarterly* 31: 434-39; October 1960.

Hatfield, Agnes B. "An Experimental Study of the Self-Concept of Student Teachers." *Journal of Educational Research* 55: 87-89; October 1961.

Heil, Louis M., and Washburne, Carleton. "Brooklyn College Research in Teacher Effectiveness." *Journal of Educational Research* 55: 347-51; May 1962.

Hughes, Marie M. "What Is Teaching? One Viewpoint." *Educational Leadership* 19: 251-59; January 1962.

Jensen, Leland E. "A Non-Additive Approach to the Measurement of Teacher Effectiveness." *Journal of Experimental Education* 30: 70-87; September 1961.

Johnston, A. Montgomery. "Factors Influencing Democratic Practices." *Childhood Education* 36: 122-25; November 1959.

Jones, Reginald L., and Gottfried, Nathan W. "Preferences and Configurations of Interest in Special Class Teaching." *Exceptional Children* 28: 371-77; March 1962.

Kleinman, Lou. "A New Dimension in Teacher Selection." *Journal of Educational Sociology* 34: 24-33; September 1960.

Lang, Gerhard. "Motives in Selecting Elementary and Secondary School Teaching." *Journal of Experimental Education* 29: 101-104; September 1960.

Medley, Donald M., and Mitzel, Harold E. "A Technique for Measuring Classroom Behavior." *Journal of Educational Psychology* 49: 86-92; April 1958.

Medley, Donald M., and Mitzel, Harold E. "Some Behavioral Correlates of Teacher Effectiveness." *Journal of Educational Psychology* 50: 239-46; December 1959.

Medley, Donald M., and Mitzel, Harold E. "Measuring Classroom Behavior by Systematic Observation." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. Chapter 6, pp. 247-328.

Mitzel, Harold E., and Dubnick, Lester. "The Relative Scholastic Ability of Prospective Teachers." *Journal of Teacher Education* 12: 73-80; March 1961.

Peck, Robert F. "Personality Patterns of Prospective Teachers." *Journal of Experimental Education* 29: 169-75; December 1960.

Popham, W. James, and Trimble, Robert R. "The Minnesota Teacher Attitude Inventory as an Index of General Teaching Competence." *Educational and Psychological Measurement* 20: 509-12; Autumn 1960.

Powell, Marvin, and Ferraro, Charles D. "Sources of Tension in Married and Single Women Teachers of Different Ages." *Journal of Educational Psychology* 51: 92-101; April 1960.

Rabinowitz, William, and Crawford, Kay E. "A Study of Teachers' Careers." *School Review* 68: 377-99; Winter 1960.

Reed, Horace B. "The Effects of Teacher Warmth." *Journal of Teacher Education* 12: 330-34; September 1961.

Ryans, David G. *Characteristics of Teachers.* Washington, D.C.: American Council on Education, 1960. 416 pp.

Ryans, David G. "Inventory Estimated Teacher Characteristics as Covariants of Observer Assessed Pupil Behavior." *Journal of Educational Psychology* 52: 91-97; April 1961. (a)

Ryans, David G. "Some Relationships Between Pupil Behavior and Certain Teacher Characteristics." *Journal of Educational Psychology* 52: 82-90; April 1961. (b)

Scandrette, Onas. "Differential Need Patterns of Women Elementary and Secondary Level Student Teachers." *Journal of Educational Research* 55: 376-79; May 1962.

Sheldon, M. Stephen. "Conditions Affecting the Fakability of Teacher-Selection Inventories." *Educational and Psychological Measurement* 19: 207-19; Summer 1959.

Smith, B. Othanel. "A Concept of Teaching." *Language and Concepts in Education.* (Edited by B. Othanel Smith and Robert H. Ennis.) Chicago: Rand McNally & Co., 1961. Chapter 6, pp. 86-101.

Tyler, Fred T. "Teachers' Personalities and Teaching Competencies." *School Review* 68: 429-49; Winter 1960.

Veldman, Donald J.; Peck, Robert F.; and McGuire, Carson. "Measuring the Value Systems of Education Professors." *Journal of Educational Psychology* 52: 330-34; December 1961.

Withall, John. "The Development of the Climate Index." *Journal of Educational Research* 45: 93-100; October 1951.

Wright, E. Muriel J. "Development of an Instrument for Studying Verbal Behaviors in a Secondary School Mathematics Classroom." *Journal of Experimental Education* 28: 103-21; December 1959.

CHAPTER VII

Curriculum and Communications Media

CURTIS PAUL RAMSEY

COMMUNICATION, curriculum, and teaching are words filled with many meanings and are yet imprecise. When effective teaching is equated to a successful communicative act, how does one know for sure that one has communicated? For that matter, how does one know for sure that one has taught any particular thing?

In the June 1961 issue of the REVIEW, Shafer discussed strategies in communications research, limited status of research methodologies and techniques, and effects of mass communication on children. The implications of communications media for curriculum are obviously great. But what research evidence exists concerning the effect of the media on curriculum? (Incidentally, almost no one seems seriously concerned about the effect that curriculum might have on communications media!)

As yet, criteria for the determination of the degree of success of transmission of meaning from one mind to another are nebulous and inadequate More is known about what contributes to failure in communication and in teaching than about what contributes to success. Theoretical proposals are far in advance of empirical research findings. Concerning this point, Carpenter (1962) warned that it is erroneous to conceive of communication processes and content as "transmissions of energy" that will lead directly to a predetermined response in the individual who has been stimulated, regardless of the nature of the mediation processes involved.

Technology of Teaching

In the June 1960 issue of the REVIEW devoted to a consideration of curriculum problems and in the April 1962 issue of the REVIEW concerned with educational media, interrelationships underlying curriculum and instructional media were noted. Finn (1960) insisted that an emerging technology of teaching must be reckoned with today. Moreover, he stressed that the implications of parallel but opposite trends in mass and individualized instruction may create a riptide of confusion in the curriculum. Of even greater danger is the possibility that the design and control of the curriculum may be taken from the hands of legitimate educational authorities through deliberate short-circuiting by foundations and other agencies.

Articles and admonitions about the role of communications media and their impact on curriculum are many. However, an entire school of media research has denied—for research purposes—that the individual learner is a mediator in the communications paradigm. Stimulus-response behaviorism, which has not addressed itself to the dynamics of individual perceptions of mass media stimulation, has instead focused myopically on a small-step, atomistic presentation and on response modules. Recently, Jordan (1963), who challenged even the Socratic premise of programed instruction, found the method far inferior to its presumed model.

Merchandising and Research

In short, there is more sound and fury in the field of communications and curriculum than there is substance. Whether one adheres to the positions taken by Finn (1960), by Ramo (1960), or by Bushnell and Cogswell (1961), the impact of discussion is not so much a million-pound thrust that can be attributed to an inexorable technological advancement as it is a million-dollar thrust built upon a capital investment.

Finn and Perrin (1962), who documented the tremendous growth of the field of programed instruction and teaching machines, found that there were 104 companies in various stages of preparation or production of devices and programs in 1962. Perhaps as few as 11 of these companies existed a decade earlier as publishers of films. Part of the enthusiasm and advocacy of new technology for teaching appears to be self-generating and self-reinforcing. Possibly to furnish some degree of control upon the runaway trend, Soles (1963) reported the development of criteria useful to distinguish between valuable educational services and commercial quackery.

However, the programed instructional devices must be viewed as more than new audiovisual aids. So potent and so well designed are the better programs that they need to be considered in a new curriculum design-construction-control perspective. Gilpin (1962) has taken the position that a truly systematic concept of instructional technology is possible; others have indicated a similar orientation.

The special significance of programed instruction is not in the conditioned response mechanism, but rather in the necessity for criterion performance designation to assess learning progress. This concept, so sharply identified in programming necessities, will force a more rigorous and better defined approach to designation of teaching tasks in other forms and other modes of instruction than have existed. Perhaps this movement will bring about the demise of those blunt and inane research designs that measure traditional teaching against some new fad or procedure whose terms are as vague and ill-defined as those used in traditional or conventional teaching. The teaching profession is surfeited with pseudo-research reports comparing total systems of instruction as if the only variable involved is a distinction between two global concepts.

In brief, there is scarce justification for optimism in the field of research related to curriculum and communication. Most project reports are selflaudatory acknowledgments of a preconceived bias toward an innovation. Many projects, which, in fact, are not research-oriented at all, constitute demonstrations—or, at best, feasibility studies—of an advocated position or procedure.

Meierhenry (1962) cautioned that, although much research is currently in progress, the prime requisite of any researcher is the ability to ask the right questions. Unfortunately, the more important questions have been neglected too long. The most important research questions to be asked at the moment seem to be related to (a) characteristics of individual learners and their approach to learning; (b) characteristics and limitations of the media, systems, and mediational combinations; (c) role of the personal teacher as a mediating influence; (d) interactional dynamics of the above three areas; and (e) development of curriculum communication theories, models, and instrumentation.

In communications theory and research related to formal curriculum purposes, it seems that investigators have proceeded scarcely beyond the witchcraft stage.

Additional Reference: Elam (1963).

Bibliography

BUSHNELL, DON D., and COGSWELL, JOHN F. "A Computer-Based Laboratory for Automation in School Systems." *AV Communication Review* 9: 173-85; July-August 1961.

CARPENTER, C. R. "Boundaries of Learning Theories and Mediators of Learning." *AV Communication Review* 10: 295-306; November-December 1962.

FINN, JAMES D. "Technology and the Instructional Process." *AV Communication Review* 8: 5-26; Winter 1960.

FINN, JAMES D., and PERRIN, DONALD G. *Teaching Machines and Programmed Learning, 1962: A Survey of the Industry.* Occasional Paper No. 3. Washington, D.C.: Technical Development Project, National Education Association, 1962. 85 pp.

GILPIN, JOHN. "Design and Evaluation of Instructional Systems." *AV Communication Review* 10: 75-84; March-April 1962.

JORDAN, JAMES A., JR. "Socratic Teaching?" *Harvard Educational Review* 33: 96-104; Winter 1963.

MEIERHENRY, WESLEY C., editor. "Needed Research in the Introduction and Use of Audiovisual Materials: A Special Report." *AV Communication Review* 10: 307-16; November-December 1962.

RAMO, SIMON. "A New Technique in Education." *Teaching Machines and Programmed Learning: A Source Book.* (Edited by Arthur A. Lumsdaine and Robert Glaser.) Washington, D.C.: Department of Audio-Visual Instruction, National Education Association, 1960. pp. 367-81.

SOLES, STANLEY. "Educational Quackery External to Our Schools: What Is a Professional Response?" *Phi Delta Kappan* 44: 299-302; March 1963.

Additional Reference

ELAM, STANLEY, issue editor. "Programed Instruction." *Phi Delta Kappan* 44: 241-302; March 1963.

CHAPTER VIII

Curriculum Development: Dynamics of Change

PAUL R. KLOHR and JACK R. FRYMIER

A CHAPTER on the dynamics of change in a review of research on curriculum planning and development might appear, offhand, to be unwarranted duplication, since each of the other chapters is also clearly concerned with change. Yet a survey of the literature of the last three years suggests the need for such a category if for no other reason than to focus attention on a crucial aspect of the total curriculum development process which will require increasing investigation in its own right if curriculum research is to mature and to take its place with other behavioral sciences.

In effect, this chapter is based on the assumption that certain dynamics of change in the curriculum development process can be identified and that research into these aspects of change will help both to describe what is happening and to predict what will happen. In previous issues of the REVIEW, this topic was subsumed under topics dealing primarily with the implementation of curriculum development, in which emphasis was placed on administrative, supervisory, and organizational arrangements.

Throughout the literature are persistent pleas for more adequate conceptual systems and improved theory. In the study of curriculum change itself, some notable beginnings were made during the period under survey. The state of the field is such that no tightly drawn categories can be used to organize the research under purview. However, the studies reviewed did suggest the following classification: search for models; diffusion of innovations; field demonstrations; redefinition of change; roles, role perception, and personality structure in change; and contributions from related fields.

Search for Models

In the quest for more adequate models for the study of curriculum change, the volume of readings in the applied behavioral sciences edited by Bennis, Benne, and Chin (1961) is a significant resource. Especially useful are the concepts the editors used to introduce each section of the volume and to relate the various sections. Taken together, these constitute a theory of planned change. This theory enables one to explore the nature of the deliberate, collaborative process which involves change agents and client systems, the two constituents of planned change. The emphasis on the deliberate and collaborative character of the relationship makes the theory especially relevant to curriculum development. In contrast to theories of change often seen as appropriate for other fields, such as business or industry, the outcome of planned change as projected by these theorists rests basically on the relationship established between the giver and the receiver of help. How well the process is understood by each and what degree of openness for examination and for possible reconstruction exists for both parties are, therefore, of central importance. Selection after selection in this volume, ranging from abstract considerations to research reports with detailed analyses of methodology, supports, in a strong, consistent way, the position taken by the editors.

Kushner and others (1962) developed a propositional inventory of hypotheses accounting for sociocultural change. Four hundred logically ordered hypotheses drawn from ethnographic studies and from statements of theory were organized into 13 categories. Each category is made up of a series of propositions illustrated by hypotheses. For example, one category holds together a number of propositions that account for the acceptance or rejection of an innovation through congruence or compatibility. Another category groups propositions that suggest techniques and motivations used by change agents to introduce innovations. Although this inventory was not seen by its authors as a theory of sociocultural change, it can serve a useful function as a general conceptual framework for examining curriculum change and for designing research to probe more deeply into this component of the total curriculum development process.

Serving a similar purpose is the work of Barton (1961a), who used a general systems approach to classify a number of studies about the behavior of organizations, including schools. His classification was made to show some principles of measurement applied to complex organizations. Concerned with the logical nature of the measurement operations that might be used to study college environments, he saw an understanding of these as a necessary condition for the designing of more effective measurements. His classification involved six categories: inputs, outputs, environment, social structure, attitudes, and activities. Although Barton asserted that he did not intend his classification to be a theory, it too suggests a general model for undertaking research into the nature of curriculum change within a school or school system.

George Maccia (1962) used von Bertalanffy's general systems theory to develop a nonrepresentational model for the study of education. The general systems theory builds on the ideas of nonlinearity and on interactions among the parts of a system. With his model, which makes central use of the concepts of input and output, Maccia generated an outline of an educational theory. Four subsystems were identified within the system: administering unit, facilitating unit, teaching unit, and inquiry unit. This model and its related theory represent a promising approach to the complex problem of assessing change in education. Moreover, they throw light on the special task of measuring and evaluating curriculum change. A similar contribution was made in the work of Elizabeth Maccia (1962), who used graph theory to develop a model and to translate it into an outline of educational theory. In effect, graph theory was used to represent characterization of the relations of persons in groups within a school setting.

Downey (1961) outlined a theoretical model of organizational behavior which makes use of the organizational equilibrium concept. An adaptation and extension of the Getzels-Guba theory of social systems, it identifies a series of procedures for change in public schools along five dimensions: cultures, politico-economic environments, institutions, individuals, and informal groups. Eight precautions to be observed in the implementation of change in a school setting were formulated.

In the quest for a more adequate sense of direction in curriculum change, Frazier (1963) identified five uses of new insights in curriculum theory and practice. His analysis was based on the 1963 Association for Supervision and Curriculum Development yearbook, which paired chapters reporting new insights from various fields with others delineating implications for curriculum. He asserted that new insights may help the educator (a) to ask new questions, (b) to re-examine assumptions, (c) to see new distinctions that need to be made, (d) to respond more completely to things already understood, and (e) to imagine new possibilities. His observation that new insights do not always reveal their most effective use unless they are studied suggests a crucial dimension in dealing with change in curriculum development operations.

Sweitzer (1962) pointed to the value of the principle of complementarity among two or more theories. Although he was concerned with theories of educational administration, the applicability of his view to curriculum theory and to the study of curriculum change is clear, whether one is seeking to identify the major questions to be investigated or trying to find a higher conceptual level than that suggested by one of the theories being compared. He demonstrated his point by an analysis of the complementarity between the nomothetic-idiographic model of Getzels and Guba and the input-output framework of Stogdill.

At this stage in the search for more adequate models with which to study the dynamics of curriculum change, one senses the need for a model of models. Sweitzer's proposal is a step in this direction.

Diffusion of Innovations

Research on the diffusion of innovations, which has increasing significance in such fields as rural sociology, has not been widely used yet in the study of curriculum change. The work of Mort and Furno (1960) stands almost alone in this regard. They proposed a sequential simplex model to account for the various classes of influence on schools. The logic of this approach rests on measures from a financial continuum taken at three points: wealth, actual expenditure, and professional salaries. Although the procedure has a sound statistical base and provides a method for comparing the impact of certain measurable influences within a group of schools, it fails, in large measure, to reveal the kinds of data required to direct curriculum change.

A broader view of diffusion research was taken by Rogers (1962) in his review and synthesis of more than 500 studies ranging from new drugs among physicians to hand tools among primitive tribes. His paradigm of the adoption of an innovation by an individual in a social setting and his method of conceptual variable analysis provide a way of relating theory to research which has applicability to the study of curriculum change. Rogers' listing of 52 generalizations in order to summarize what is known about the diffusion of innovations suggests many specific studies yet to be made. Studies might be undertaken in a school situation where curriculum change is under way to test, for example, his proposition that personal influence from peers is most important at the evaluation stage in the adoption process and less important at other stages. With data from this kind of diffusion of innovation research, the Mort instrument might become more useful.

Although not concerned so directly with the problems of diffusion of innovations, several investigators have thrown light on closely related aspects. Mackenzie (1962) and Huebner (1962) examined the political qualities of curriculum change. Mackenzie proposed a model to show the interrelations of three factors in change: social context, participants, and participant interaction. He placed these on a vertical axis and located three stages of outcomes in the curriculum development process on a horizontal axis: identification of focus, development of proposals, and implementation through teaching. He then stated six propositions about curriculum innovations to test the model and called for the development and testing of other models to the end that increased ability to explain and control the process of curriculum change might be fostered. Huebner urged that educators use the concepts of political science to help analyze the field of curriculum, redefine their roles, and direct new inquiry. As illustrations, he examined action, elites, and ideologies as these concepts relate to educational operations.

An analysis of the sociology of reading research made by Barton (1961b) has implications for the study of curriculum change in general. He charted the relationships among status groups and institutions that influence reading research and developed a scheme that permits an examination of the social organization of such research, the methods used, the channels by which it is interpreted and communicated, and the response of educators to it. This analysis in a specific field lends striking support to the more general thesis advanced by Mackenzie and Huebner and suggests a framework for investigating the nature of diffusion in many other areas of education, including curriculum innovation.

In the field of education, in curriculum change, as well as in such other areas as leadership theory and research, there appears to be little concern with the problem of power and power relationships. The Mackenzie, Huebner, and Barton efforts are exceptions. Political scientist Janda

(1960) called for a conception of leadership explicated in terms of power. Gross (1962), a sociologist, described the lines of integration around which action is centralized and maintained in a social organization such as a school and delineated eight bases of power undergirding these lines. He concluded that an individual must have a clear understanding of the power available to him and to others in undertaking change.

The research and theory-building efforts that focus on the problems of diffusion of innovation in education, especially innovations in curriculum, make one need clear—the need for more adequate agencies for change. Numerous proposals have been advanced. The most viable, among those surveyed, is that made by Clark (1962), who proposed the establishment of new mechanisms for the development of educational research. He called for the establishment of a series of research institutes, which, in addition to providing support for basic and applied research in six categories including curriculum, would also provide for the needed field testing of research. He recommended also the establishment of state and regional experiment centers for demonstration and dissemination of research findings to practitioners. These proposals have major implications for changing curriculum.

Field Demonstrations

Certain efforts to study the dynamics of curriculum change tend to fall into a field demonstrations category. These efforts have many characteristics of descriptive research.

An illustration is the study undertaken in New York for the State Education Department. With a grant from the Fund for the Advancement of Education of the Ford Foundation, Brickell (1961) identified barriers related to the promotion of change in selected schools. He visited 30 school systems for an average of nine hours each and spent seven days in New York City schools. Thirteen teacher education institutions were visited, and personnel in the State Education Department were interviewed. A series of 30 generalizations about the process of change within local schools and about the role of outside agencies make up the body of the report. Brickell concluded that there are three phases of instructional innovation—design, evaluation, and dissemination—and that the most serious block to instructional improvement is a failure to distinguish among these three. In his view, observations of the process of change in medicine, agriculture, and industry supported these conclusions. As a solution to the problem he identified, Brickell proposed the creation of an agency to stimulate and to finance the design and evaluation of new instructional programs. To disseminate the programs developed in this way, he recommended the creation of regional development units.

Two serious weaknesses, which tend to characterize most descriptive studies of change, stand out in this study. The first is its narrow focus. Brickell grossly oversimplifies the complex nature of change by equating it with shifts in the arrangement of six structural elements of the institution: teachers, subjects, students, methods, times, and places. Second, there is an absence of any clear conceptual framework to provide the necessary guidelines for assessing the quality of the observed change. These weaknesses limit markedly the usefulness of the study by closing off from wider public consideration the basic data on which the specific recommendations for New York State rest. Its transfer value, or applicability, in other situations is thereby seriously reduced.

Trump and Baynham (1961) reported, for the Commission on the Experimental Study of the Utilization of the Staff in the Secondary School, a study which focused on change. It described the projects developed in 100 junior and senior high schools involved in the Commission's study of staff utilization over approximately a four-year period. The authors did not include statistical data collected by the schools, inasmuch as some of these have been made available elsewhere. However, they did generalize on such data and concluded, among other things, the following: Innovations in the schools provided students with improved learning opportunities. Teachers were more satisfied. Principals exerted more effective instructional leadership. General interest in experimentation was created in school and community. They then addressed themselves to the question of how other schools should organize for change. Among the necessary conditions they recognized was the human factor: Change involves people. They also emphasized the need for sound research and evaluation procedures.

This study suffers from shortcomings similar to those of the Brickell report. Despite its insistence that individuals and groups are central in the process of curriculum change and that adequate research and evaluation procedures must be used, great voids exist between the available data and the conclusions reported; that is, one is required to accept their conclusions almost on faith. Herrick commented in the June 1960 REVIEW that the extensive surveys of the Commission during the period 1957-60 were centered on the administrative and mechanical aspects of staff reorganization rather than on "corollary matters of curriculum change." This view applies to the more recent work of the Commission as well. Yet, because of the generally confused state of curriculum development research, especially that which deals with the dynamics of change, and because of the sponsorship of the Brickell and the Trump and Baynham reports, which gives them high public visibility, both studies undoubtedly will wield considerable influence.

Frazier (1962) raised some basic questions about the role of "administrators as curriculum leaders," which he applied to Brickell's work. They have relevance also for many aspects of the Trump report as well as for a number of other efforts to analyze the dynamics of curriculum development using a field demonstration approach. Frazier was concerned, for example, that Brickell excluded classroom practices that involved the teacher at work with his own pupils. In fact, the crucial role of the classroom teacher in curriculum change seems to be denied. Frazier urged that the process of curriculum development be redefined to make room for new functions of both teacher and administrator and that purposes be kept sharply in review throughout the process.

Redefinition of Change

Taba (1962) advanced a strategy for changing the curriculum which takes into account the questions raised by Frazier. Moreover, pursued as an approach to curriculum experimentation, it could provide the data that seem to be missing in the Brickell and Trump and Baynham analyses of change. She called for a deliberate inversion of the common procedure of starting with a general design. Instead, she proposed that innovation start at the teaching level with the planning of specific teaching-learning units. The results of experimentation at this level then would provide a basis for the creation of a general design during a later phase of the total process. This approach is calculated to infuse theory into the operations of the practitioner from the outset. Although Taba's proposal is speculative, she drew on a wealth of experience as a curriculum consultant and marshaled substantial support for her position from the behavioral sciences. Carefully planned experimentation with the inversion approach would add significant knowledge to the curriculum development field and would represent an important step forward in the redefinition of curriculum change in order to make it researchable.

Within the descriptive category, the work of McNally and others (1960) stands in sharp contrast to the Brickell report in terms of its potential use in assessing curriculum change or in planning innovations. This volume analyzed organized efforts to improve curriculum in seven cases ranging from a relatively small suburban school system to a state-wide program. From this analysis, the authors formulated 11 guidelines for improving programs of instruction and proposed criteria for evaluating change in program. Both formulations make explicit the values involved in the curriculum development process. Thus, they furnish essential resources for curriculum theory and practice not provided in many descriptive studies that fail to recognize the role that values have in the total process.

Another example of descriptive research that has promise for contributing to the redefinition of curriculum change is the study Guggenheim (1961) made of 16 school districts in New York State to identify patterns of staff utilization. Using a case study approach to appraise the communications systems in the schools, he followed a unit of information that was involved in a referral on a deviant pupil through the organizational system and found out from each recipient of communicated information data how information was received. The Lazarsfeld-Robinson pseudotest approach to quantification was used. Guggenheim found that the full utilization of staff is highly related to the kind and quality of a school's channels of communication. This suggests that any evaluation of curriculum change well might include a study of the relation of communications to staff utilization and school organization. Moreover, the research design used in this exploratory study can be adapted well to the larger task.

Teckman (1961) studied the influence of state departments of education and regional accrediting associations on experimental projects in 39 selected secondary schools. Using an extensive questionnaire and an analysis of standards, policy statements, and published criteria, he found that few local school leaders had received support for experimentation from these sources and that the standards and policies of such agencies did not encourage local schools to conduct experimentation. Moreover, the majority of leaders sampled in the study indicated that the present standards of state departments and regional associations are unable to keep pace with postulated new practices in secondary education. He called for a redefinition of the role of these agencies as change agents.

Alexander (1962a, b) identified some priorities in curriculum development including the establishment of a system of curriculum research councils at the local, state, and regional or national levels. He sought in this way to foster a more adequate screening, review, and dissemination of curriculum research.

Wrenn (1962) speculated about the conditions that are likely to keep changes which were recommended by the American Personnel and Guidance Association's Commission on Guidance in American Schools from taking place and identified three cultural encapsulations that contribute to resistance to change and four preventive measures that might be taken to overcome this resistance. Although aimed at the counselor, his recommendations may be generalized to other educational personnel who have helping or change-agent responsibilities similar to those of the counselor. Seen in this light, they represent an effort to redefine the nature of change.

Roles, Role Perception, and Personality Structure in Change

In moving from the more speculative theory-building kinds of operations within the total spectrum of research on dynamics of curriculum change, certain descriptive studies have been noted. Some of them are theory-oriented, while others function as field demonstrations with little, if any, relation to a larger conceptual framework. However, the largest number of investigations made during the period under survey falls within yet another category. These are the specific, often isolated, studies that probe a single aspect of the curriculum change complex. The most prevalent focus for this research has been on roles, role perceptions, and personality structure.

Perceptions of changes in key instructional practices held by 804 elementary school and 918 secondary school principals were identified through a questionnaire survey and reported in a monograph by the National Education Association, Project on Instruction (1962). To identify some of the forces influencing change, principals rated a list of 14 possible sources of influence. Of these, the respondents indicated that local school officials had exerted the greatest influence. Faculty influence was rated second, and state school officials were third, with research studies holding fourth place. Various national studies in academic subjects were a close fifth on the composite list. In general, the replies indicated that national studies have increased significantly in their influence from 1956 to 1961 but that little additional increase is expected between 1961 and 1966.

Anderson and Anderson's (1962) study of more than 10,000 children and teachers from eight countries with projective, incomplete stories indicated clearly that authoritarian cultures do tend to induce defensive behaviors among children and teachers more frequently than do democratic cultures. One can infer from these studies that the extent and nature of curriculum change may be directly related to the kinds of personality characteristics evident among the teachers involved. Future studies should be initiated to explore this implication.

Flanders' (1960) intensive investigation of teacher influence, pupil attitudes, and achievement is an especially important study in that it was carefully framed within explicit theory and was aimed at developing instruments and procedures designed to help teachers change their classroom behaviors. Moreover, it included, in the research design, a liaison committee of parents, teachers, principals, and supervisors to oversee the entire research effort. Interaction analysis, a precise technique for studying verbal activities in a classroom situation, includes working with teachers so that professed goals can be compared to attained goals. The fact that Flanders translated the results of this research into a series of filmstrips and recordings for use with teachers in in-service education programs is also noteworthy. Few researchers attempt to bridge the gap between their activities and classroom practice. These materials are especially significant in that they provide a way to induce curriculum change by involving teachers directly in a process which has been substantiated carefully by research. Few in-service programs have such a base.

In the area of interpersonal perception, Rosen, Levinger, and Lippitt (1960) used the individual interview to study 64 preadolescent, underprivileged boys in a four-week camp situation. They found that, in a peer group, the desire for change in oneself or in others is related to the relative resourcefulness of individuals. While this finding clearly cannot be generalized to groups of teachers and administrators involved in curriculum development, it does suggest possibilities for a number of other inquiries into the whole matter of how the concept of resourcefulness functions in the way individuals and groups see themselves in the process of change.

Schutz (1961) grouped 70 educational leaders, involved in a series of 10 meetings, according to their reactions on a questionnaire designed to measure an individual's expressed and wanted behavior in the interpersonal areas of inclusion, control, and affection. After meeting six times, these groups composed of persons with individual interpersonal needs as determined by the questionnaire were able to identify accurately their own group description from among five possibilities. This research suggests strongly, as does that of Rosen, Levinger, and Lippitt, that certain group tasks might be undertaken more effectively by individuals grouped in terms of a measure of their interpersonal needs.

In another research undertaking, Schutz and others (1961) applied the FIRO theory to a study of the interpersonal behavior of school administrators. Eighty scales were developed in which the Guttman scaling technique was used. The findings throw light on the nature of individuals who commonly fulfill the change-agent role in the curriculum development process. For example, teachers were found to respect the principal similar to themselves with regard to how he handles difficult situations. Also identified in the study were two groups of people whose attitudes toward education seemed closely related to the way in which an individual perceives others. One group displayed positive attitudes toward others and viewed the functioning of the school in a more open manner. The other group displayed a more conventional orientation and viewed the total educational process within a less flexible framework.

Seeman (1960) used change as one of four aspects of leadership style in the total cultural approach he took to the situational study of 77 school executives in 26 Ohio communities. His findings have implications for identification and training of leaders to guide curriculum development. For example, he discovered that there is a clear carry-over of status ratings and orientations from one situation to another. This finding suggests a number of related questions to be investigated, if one assumes that curriculum change involves the community as well as the status leaders in education.

Comparing teachers' and administrators' perceptions of the teacher's role, Fishburn (1962) noted a fundamental disagreement between the two groups. Whereas teachers saw their most important functions as mediator of the culture and director of learning, administrators saw liaison between school and community as the teacher's chief role. Hencley's (1961) study of board members' and superintendents' perceptions of their own and each others' roles also illustrates that these groups have conflicting perceptions about their areas of responsibility and authority. Since Scarborough and Harris (1962) reported that graduate students' perceptions of leadership roles corresponded directly to those of practicing administrators and consultants, it may be that programs of training for educational leaders are helping perpetuate certain basic misconceptions among some segments of the profession.

Drawing on his experiences in psychotherapy, Rogers (1961) outlined a stasis-process continuum of personality change. Exploring the direct relationships between these ideas and education, he then spelled out the conditions that seem to be present when an individual moves from stasis toward process in the act of becoming. Assuming that motivation for change and learning springs from life itself, Rogers describes the conditions that can be provided by the therapist and that seem essential for desirable change. These include congruence, unconditional positive regard, empathic understanding, and awareness on the part of the client that these conditions exist in the therapist. Although the concepts have not been substantiated by extensive research in either psychotherapy or education, the framework of the theory should become clearer and have more relevance as hypotheses regarding change emerge and are tested within its confines.

Bills (1963) cited several studies based directly on Rogers' concepts. The more a teacher was open to his experience, the more his students regarded him as congruent, empathic, positive, and unconditional. The locus of decision making for the more-process teachers was centered in the class, including the teacher. On the other hand, the locus of decision making for the less-process teachers was seen as centered either in the teacher (autocratic) or in the students (laissez-faire). In another study cited, Bills reported that teachers and administrators who were open to experience changed significantly in their attitudes towards themselves and others as a result of a summer workshop experience. Those persons who were less open to experience, however, were unaffected by the workshop. These studies suggest that those who seek to change the curriculum by involving teachers must contrive in-service education efforts that are tailored to meet the degree of openness of the various teachers involved. No single approach to the problem seems adequate.

These studies, taken as a sample of wide-ranging investigations into the dynamics of role, role perceptions, and personality structure in educational change, suggest the importance of continued investigation in curriculum development programs of the general sensitivity training approach which has characterized efforts in the field of business research. Helfant's (1961) study of the effects of three human relations workshops held in the Manhasset Public Schools in New York underscores this need. He found that group therapy was effective in modifying the attitudes of administrative personnel and in maintaining a preventive mental hygiene climate.

Shumsky and Mukerji (1962) studied the role of the consultant in bringing about changed behavior in action research projects and concluded that (a) the way the consultant relates himself to the teacher-researcher during the implementation phase of the undertaking is highly significant and (b) the deepening of the teacher's self-understanding is a key factor in changed behavior.

In a study of the personality characteristics of teachers who were resistant to change, Myers and Torrance (1961) identified the following characteristics: authoritarianism, defensiveness, domination by time, insensitivity to pupils' needs, low energy level, preoccupation with information-giving functions, intellectual inertness, disinterest in promoting initiative in pupils, preoccupation with discipline, and unwillingness to give themselves to the teaching-learning compact. These findings were based on an analysis of the responses of 114 teachers in 14 states to the request that they report how they applied five principles in their teaching. The principles, which suggested ways of rewarding creative thinking in the classroom, were (a) treat questions with respect; (b) treat imaginative ideas with respect; (c) show your pupils that their ideas have value; (d) occasionally have pupils do something for practice without the threat of evaluation; and (e) tie in evaluation with causes and consequences.

Cox (1961) reviewed four studies made at the University of Georgia, the findings of which strongly supported the importance of viewing the instructional supervisor's role from various perspectives. A similar outcome was reported by Guss (1961), who chaired an Association for Supervision and Curriculum Development research committee's survey of the perceptions of the supervisor which various individuals held.

Contributions from Related Fields

Recent summaries of research and theory in the field of business management suggest approaches that merit testing in any comprehensive effort to study the nature of curriculum change. Likert (1961) reviewed extensive research carried out since 1947 at the Institute for Social Research of the University of Michigan and developed, among other things, (a) principles that describe the interaction-influence system in an organization and (b) a comparative analysis of organizational and performance characteristics of different management systems, ranging from authoritative to participative.

Tannenbaum, Weschler, and Massarik (1961) reported work representative of a 10-year research and theory building effort of the Human Relations Research Group of the Graduate School of Business Administration, University of California, Los Angeles. A basic idea underlying the work of this group was that leadership, influence, and organizational effectiveness all bear directly on improved human capacities to diagnose interpersonal situations and to act effectively on the basis of this diagnosis. In short, their view of change, which builds on the pioneer concepts of Lewin, Rogers, Moreno, and Bion, is reflected in the more recent efforts of such men as Benne, Bradford, Lippitt, and Thelen.

Vroom (1960), working with data obtained from 1,676 employees of a large industrial organization, found that persons tend to attribute their own attitudes, opinions, and goals to persons, groups, and organizations toward which they have a positive attitude and to deny them in persons, groups, and organizations toward which they have a negative attitude.

In a comprehensive effort at theory building, Presthus (1962) analyzed research from several social science areas as it applied to the behavior of individuals within bureaucratic organizations. Having underscored the dynamics that create conflict between persons who strive for professional competence and those who seek status and power through administrative roles, he then delineated the aspects of large organizations that affect the functional relationships of members within the structure. Individuals who identify strongly with the organization, who aspire to administrative posts, and whose major satisfactions come from acceptance of and involvement with the institutional values and goals are characterized as upward mobiles. Those whose patterns of accommodation include simultaneous rejection of the organization's purposes and acceptance of the security available for those who go along are identified as the indifferents. Neither of these major groups promotes or encourages critical examination of an organization's structure, purposes, or values. Ambivalents, however, who represent a dissident minority group whose behaviors are essentially intellectual and introspective, provide concern for innovation and change through their probing criticism and resistance to authority. Seldom functioning even in relative harmony with the organization, the ambivalents ultimately alienate themselves from the organizational elite. This model should generate fruitful hypotheses for researchers interested in the impact of institutional structure upon educational goals and the nature of change within the institutional setting.

Need for Further Research

By way of conclusion, perhaps the best way to sense the state of the field—beyond whatever insight one gains from the preceding review of the literature dealing with the dynamics of curriculum change—is to focus the several major problems that clearly require further research. Three such problems are apparent.

One of the basic questions that has not been studied directly is whether involvement in the research process itself is an effective way to bring about curriculum change. Barnes and others (1960) published a descriptive handbook of seven classroom teachers' efforts to improve instruction through research. It implied that those problems that stem directly from classroom teachers' interests and difficulties offer unique possibilities for curriculum change by involving those most immediately concerned in practical studies. This is, essentially, the action-research approach.

Inherent in this approach is the basic idea that curriculum research from the bottom up is the most productive and most effective way to bring about curriculum change. A few school systems, however, maintain highly organized research divisions whose purpose is to initiate curriculum research from the top down. Which approach is more effective in encouraging significant curriculum change should be examined. Speculative treatments of this problem are necessary but not sufficient.

Two other problems are related to the question, Who should initiate curriculum research? The first concerns research design. Many persons in university and governmental research centers are urging educational researchers to devise more elaborate and complex research designs. In an obvious effort to cope with the many variables inherent in the total process of educational change, researchers are encouraged to conceptualize their projects in such a way that as many of the important variables as possible are brought within the framework of their design in some manipulatable or controllable way.

These efforts naturally result in research reports that are complex and difficult to understand. Some reports regularly employ statistical concepts and procedures that the typical classroom teacher or school administrator has not been trained to understand. In view of the resulting maze of detail and unfamiliar terminology, there may actually be a widening gap between those producing research data and those most interested in utilizing the results of these studies in classroom situations.

Two hypotheses seem worth exploring in this regard. First, Is there a need in the field of education for some type of dissemination service or agency that would function in a manner comparable to the county agent in agriculture to help teachers and administrators relate the best research results available to the local school situation? Could such an agency both strengthen and broaden the communication bridge between researchers and practitioners to the extent that curriculum would be affected in any appreciable way? Brickell (1961) and Clark (1962) both made proposals along this line.

Second, Should those persons in leadership positions work directly to establish policies and structures, within their own school districts, which regularly engage teachers, parents, and administrators in efforts that cut across organizational and grade level lines in order to improve instruction through research at the local level? McNally and others (1960) and Alexander (1962b) outlined what appear to be viable approaches to this problem, but further large-scale research is required.

Although there is apparently no research at all which bears directly on one final problem, the growing surge of interest in both professionalism and curriculum concurrently would seem to suggest that educators might well explore the impact of the structure of the profession itself on curriculum development and curriculum change. Does the nature of the professional relationships that currently prevail among principals, teachers, and superintendents inhibit the process of curriculum development in any important way? In other words, Is there any evidence that the general reluctance of teachers to initiate or participate in extensive curriculum change is a direct function of the kinds of relationships they have with their administrative superiors? Can those who work directly with curriculum bring about significant improvement in the instructional program in a profession dominated and controlled by persons whose interests in the curriculum are peripheral?

Though no direct evidence is available, it would seem that those groups that have strengthened their professional status through stronger organization and that impose definite limitations upon their membership through internal controls have simultaneously increased their effectiveness. There would seem to be more than a casual relationship between the rise of the physicians in professional stature, for instance, and their ability to achieve the basic purpose of maintaining ever higher levels of personal health among the public at large. In practice, this increased effectiveness seems to have come about by the physicians' efforts to contrive ways of helping themselves to be both aware of and affected by the most recent medical research. Significant progress in curriculum development is not possible until there is a general strengthening of the teaching profession through shifts in the power structure. Clearly, research into the dynamics of curriculum change must cope with this kind of problem, however complex it is.

Bibliography

ALEXANDER, WILLIAM M. "Assessing Curriculum Proposals." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 11-21. (a)

ALEXANDER, WILLIAM M. "Some Priorities in Curriculum Development." *Theory into Practice* 1: 222-27; October 1962. (b)

ANDERSON, HAROLD H., and ANDERSON, GLADYS L. "Social Values of Teachers in Rio de Janeiro, Mexico City, and Los Angeles County, California: A Comparative Study of Teachers and Children." *Journal of Social Psychology* 58: 207-26; December 1962.

BARNES, FRED P., and OTHERS. *Practical Research Projects.* Springfield: Illinois Association for Supervision and Curriculum Development (100 E. Edwards St.), 1960. 56 pp.

BARTON, ALLEN H. *Organizational Measurement and Its Bearing on the Study of College Environments.* New York: College Entrance Examination Board, 1961. 82 pp. (a)

BARTON, ALLEN H. "The Sociology of Reading Research." *Teachers College Record* 63: 94-101; November 1961. (b)

BENNIS, WARREN G.; BENNE, KENNETH D.; and CHIN, ROBERT, editors. *The Planning of Change: Readings in the Applied Behavioral Sciences.* New York: Holt, Rinehart and Winston, 1961. 781 pp.

BILLS, ROBERT E. "Education Is Human Relations." *New Insights and the Curriculum.* 1963 Yearbook. Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1963. Chapter 8, pp. 165-87.

BRICKELL, HENRY M. *Organizing New York State for Educational Change.* Albany: New York State Department of Education, 1961. 107 pp.

CLARK, DAVID L. "Educational Research: A National Perspective." *Educational Research: New Perspectives.* (Edited by Jack A. Culbertson and Stephen P. Hencley.) Danville, Ill.: Interstate Printers and Publishers, 1963. Chapter 1, pp. 7-17.

COX, JOHNNYE V. "The Supervisor at Work." *Educational Leadership* 19: 131-38; November 1961.

DOWNEY, LAWRENCE W. "Organizational Theory as a Guide to Educational Change." *Educational Theory* 11: 38-44; January 1961.

FISHBURN, C. E. "Teacher Role Perception in the Secondary School." *Journal of Teacher Education* 13: 55-59; March 1962.

FLANDERS, NED A. *Teacher Influence, Pupil Attitudes, and Achievement.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Program, Project No. 397. Minneapolis: University of Minnesota, November 30, 1960. 121 pp. (Not available)

FRAZIER, ALEXANDER. "The Superintendent: His Own Curriculum Director?" *Educational Leadership* 20: 38-43; October 1962.

Frazier, Alexander. "New Insights Toward What Ends?" *New Insights and the Curriculum.* 1963 Yearbook. Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1963. pp. 1-8.

Gross, Llewellyn. "A Theory of Power and Organizational Process." *School Review* 70: 149-62; Summer 1962.

Guggenheim, Fred. "Application of Communication Analysis to Studying Staff Utilization." *Journal of Educational Sociology* 35: 79-90; October 1961.

Guss, Carolyn. "How Is Supervision Perceived?" *Educational Leadership* 19: 99-102; November 1961.

Helfant, Kenneth. "Group Psychotherapy with Teachers and Administrators in a Public School." *Educational Leadership* 19: 187-90; December 1961.

Hencley, Stephen P. "The School Superintendent and His Role: A Conflict Typology." *Educational Research Bulletin* 40: 57-67; March 1961.

Huebner, Dwayne. "Politics and the Curriculum." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 87-95.

Janda, Kenneth F. "Towards the Explication of the Concept of Leadership in Terms of the Concept of Power." *Human Relations* 13: 345-63; November 1960.

Kushner, Gilbert, and others. *What Accounts for Sociocultural Change?* Chapel Hill: Institute for Research in Social Science, University of North Carolina, 1962. 51 pp.

Likert, Rensis. *New Patterns of Management.* New York: McGraw-Hill Book Co., 1961. 279 pp.

Maccia, Elizabeth Steiner. *An Educational Theory Model: Graph Theory.* Occasional Paper 62-125. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 38 pp. (Mimeo.)

Maccia, George S. *An Educational Theory Model: General Systems Theory.* Occasional Paper 62-126. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 34 pp. (Mimeo.)

Mackenzie, Gordon N. "Politics of Curriculum Change." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 76-86.

McNally, Harold J., and others. *Improving the Quality of Public School Programs.* New York: Bureau of Publications, Teachers College, Columbia University, 1960. 331 pp.

Mort, Paul R., and Furno, Orlando F. *Theory and Synthesis of a Sequential Complex.* New York: Institute of Administrative Research, Teachers College, Columbia University, 1960. 104 pp.

Myers, R. E., and Torrance, E. Paul. "Can Teachers Encourage Creative Thinking?" *Educational Leadership* 19: 156-59; December 1961.

National Education Association, Project on Instruction. *The Principals Look at the Schools: A Status Study of Selected Instructional Practices.* Washington, D.C.: the Association, 1962. 75 pp.

Presthus, Robert. *The Organizational Society.* New York: Alfred A. Knopf, 1962. 323 pp.

Rogers, Carl R. *On Becoming a Person.* Boston: Houghton Mifflin Co., 1961. 420 pp.

Rogers, Everett M. *Diffusion of Innovations.* New York: Free Press of Glencoe, 1962. 367 pp.

Rosen, Sidney; Levinger, George; and Lippitt, Ronald. "Desired Change in Self and Others as a Function of Resource Ownership." *Human Relations* 13: 187-93; August 1960.

Scarborough, Clive, and Harris, Ben M. "Graduate Students' Perceptions of Leadership Roles." *Journal of Teacher Education* 13: 60-64; March 1962.

Schutz, William C. "On Group Composition." *Journal of Abnormal and Social Psychology* 62: 275-81; March 1961.

Schutz, William C., and others. *Procedures for Identifying Persons with Potential for Public School Administrative Positions.* Berkeley: University of California, 1961. 192 pp.

Seeman, Melvin. *Social Status and Leadership: The Case of the School Executive.* Educational Research Monograph No. 35. Columbus: Bureau of Educational Research and Service, Ohio State University, 1960. 156 pp.

SHUMSKY, ABRAHAM, and MUKERJI, ROSE. "From Research Idea to Classroom Practice." *Elementary School Journal* 63: 83-86; November 1962.
SWEITZER, ROBERT E. "An Assessment of Selected Theories of Administration." *Educational Research: New Perspectives.* (Edited by Jack A. Culbertson and Stephen P. Hencley.) Danville, Ill.: Interstate Printers and Publishers, 1963. Chapter 13, pp. 199-231.
TABA, HILDA. *Curriculum Development: Theory and Practice.* New York: Harcourt, Brace and World, 1962. 529 pp.
TANNENBAUM, ROBERT; WESCHLER, IRVING R.; and MASSARIK, FRED. *Leadership and Organization: A Behavioral Science Approach.* New York: McGraw-Hill Book Co., 1961. 456 pp.
TECKMAN, CHARLES E. *The Influence of State Departments and Regional Accrediting Associations on Secondary-School Experimentation.* Doctor's thesis. Columbus: Ohio State University, 1961. 240 pp.
TRUMP, J. LLOYD, and BAYNHAM, DORSEY. *Focus on Change: Guide to Better Schools.* Chicago: Rand McNally & Co., 1961. 147 pp.
VROOM, VICTOR H. "The Effects of Attitudes on Perception of Organizational Goals." *Human Relations* 13: 229-40; August 1960.
WRENN, C. GILBERT. "The Culturally Encapsulated Counselor." *Harvard Educational Review* 32: 444-49; Fall 1962.

Additional References

ARGYRIS, CHRIS. *Interpersonal Competence and Organizational Effectiveness.* Homewood, Ill.: Irwin-Dorsey Press, 1962. 292 pp.
COSTIN, FRANK. "Attitudinal Outcomes of Child Psychology Courses Having Different Orientations." *Journal of Psychology* 51: 113-19; January 1961.
DODSON, DAN W. "Power as a Dimension of Education." *Journal of Educational Sociology* 35: 203-15; January 1962.
DOWNEY, LAWRENCE. "Direction amid Change." *Phi Delta Kappan* 42: 186-91; February 1961.
DURRELL, DONALD D. "Implementing and Evaluating Pupil-Team Learning Plans." *Journal of Educational Sociology* 34: 360-65; April 1961.
FOX, DAVID J., and LORGE, IRVING. "The Relative Quality of Decisions Written by Individuals and by Groups as the Available Time for Problem Solving Is Increased." *Journal of Social Psychology* 57: 227-42; June 1962.
GALLO, PHILIP S., and MC CLINTOCK, CHARLES G. "Behavioral, Attitudinal, and Perceptual Differences Between Leaders and Non-Leaders in Situations of Group Support and Non-Support." *Journal of Social Psychology* 56: 121-33; February 1962.
GOLDMAN, MORTON; BOLEN, MERLYN E.; and MARTIN, RANDALL B. "Some Conditions Under Which Groups Operate and How This Affects Their Performance." *Journal of Social Psychology* 54: 47-56; June 1961.
GOLEMBIEWSKI, ROBERT T. *The Small Group: An Analysis of Research and Operations.* Chicago: University of Chicago Press, 1962. 303 pp.
GOWIN, D. B.; NEWSOME, G. L.; and CHANDLER, K. A. "A Scale To Study Logical Consistency of Ideas About Education." *Journal of Psychology* 51: 443-55; April 1961.
HARE, A. PAUL. *Handbook of Small Group Research.* Glencoe, Ill.: Free Press, 1961. 512 pp.
HOOD, WILLIAM ROBERT, and SHERIF, MUZAFER. "Verbal Report and Judgment of an Unstructured Stimulus." *Journal of Psychology* 54: 121-30; July 1962.
JOSHI, VIDYA. "Attitude Towards Reception of Technology." *Journal of Social Psychology* 58: 3-7; October 1962.
KOWITZ, GERALD T. "The Change and Improvement of School Practices." *Phi Delta Kappan* 42: 216-19; February 1961.
LANZETTA, JOHN T., and ROBY, THORNTON B. "The Relationship Between Certain Group Process Variables and Group Problem-Solving Efficiency." *Journal of Social Psychology* 52: 135-48; August 1960.
MANN, J. W. "Group Relations in Hierarchies." *Journal of Social Psychology* 54: 283-314; August 1961.

MANN, JOHN H., and MANN, CAROLA HONROTH. "The Relative Effectiveness of Role Playing and Task Oriented Group Experience in Producing Personality and Behavior Change." *Journal of Social Psychology* 51: 313-17; May 1960.

MITCHELL, JAMES V., JR. "A Re-Analysis of the Data in Kuenzli's Study on 'Preferences for High and Low Structure Among Prospective Teachers.'" *Journal of Social Psychology* 54: 269-72; August 1961.

SHAW, MARVIN E. "A Serial Position Effect in Social Influence on Group Decisions." *Journal of Social Psychology* 54: 83-91; June 1961.

SNOW, C. P. *Science and Government.* Cambridge, Mass.: Harvard University Press, 1961. 88 pp.

TAYLOR, MARVIN; CROOK, ROBIN; and DROPKIN, STANLEY. "Assessing Emerging Leadership Behavior in Small Group Discussion Groups." *Journal of Educational Psychology* 52: 12-18; February 1961.

ZAGONA, SALVATORE V., and MACKINNON, WILLIAM J. "Open-Mindedness and Hypotheses of Interdependence Between Libertarian and Equalitarian Processes." *Journal of Psychology* 52: 347-61; October 1961.

CHAPTER IX

Curriculum Research: Problems, Techniques, and Prospects

JAMES B. MACDONALD and JAMES D. RATHS

THIS CHAPTER is concerned with research on curriculum research. In other words, its purpose is to locate studies that have focused on models, procedures, designs, and techniques in curriculum research, which enlarge understanding of the task and provide guidelines for curriculum research workers.

Although reports of specific studies designed to attack this problem area are sparse indeed, discussion of the problem is in evidence. Under the advice of the chairman of this issue, the authors have attempted to put together a commentary on research in curriculum research which, it is hoped, will stimulate future studies in this area.

What is research on curriculum research? If one may take the lead from Gage (1963) concerning research on teaching, four concept definitions seem appropriate: curriculum, research, curriculum research, and research on curriculum research.

Curriculum is defined as status of or change in pupil behavior which is pertinent to the goals of the school. Research is defined, after Gage (1963), as activity aimed at increasing one's power to understand, predict, and control events of a given kind. It follows that curriculum research is research in which one or more of the variables represent the status of or change in pupil behavior which is pertinent to the goals of the school. Research on curriculum research, then, becomes an activity aimed at increasing the power of research methods designed for understanding, predicting, and controlling curriculum phenomena.

Problem of Relational Contexts

Research on curriculum research can profitably look at the problem of relational contexts. There are at least three contexts which come to mind. First is the context of theory. As Coladarci (1960), Gage (1963), Getzels and Guba (1957), Herrick (1961), and Willower (1962) have justly noted, research in curriculum must be guided by use of models, theories, or paradigms if hypotheses that can be related meaningfully to a body of knowledge are to be generated.

Much curriculum research has been unrelated or unrelatable to theoretical frameworks. The resulting dilemma has been both a lack of coherent direction in curriculum research and a failure to provide adequate frames of reference for assessing progress in the field or for correlating research efforts. Lacking the conceptual tools of older sciences, curriculum research must combine the process of hypothesis construction with theory development.

The significance of hypotheses without theories has been criticized. For example, Carroll (1961) suggested that educational psychologists have failed to provide sufficient knowledge about some central problems in school learning. He stressed that research should result in statements about what effects teaching procedures and other actions would have upon the kinds of student behavior one does or does not desire as outcomes of education. Fisher and Roth (1961) suggested multiple hypotheses, for any study, as a way of examining relevant phenomena in their complexity. Multiple hypotheses and their stated interrelationships would, by necessity, involve the use of models, theories, or paradigms.

Joncich (1962) attempted to describe the history of the development of educational research efforts and to indicate some of the present problems. Central to her criticism was a concern that curriculum research had been, and was still, guided by the progressives rather than by the neutral and objective workers in the field. Apparently, her statement refers to the substitution of a philosophy for a theory, as a generator of hypotheses. Her caution is well put. However, the authors are much more concerned about biased research, which is done to achieve some practical goal in school systems, than they are about the danger of progressivism in research. It would be useful, however, to have empirical evidence of the values which researchers hold in relation to their research activities.

It would appear that curriculum research efforts could be improved considerably through the use of adequate frames of reference within which hypotheses are generated. The testing of these hypotheses would have the dual effect of providing data both about theory and about the original concern or felt need. In addition, it is only as this criterion is met that null results are able to contribute in any meaningful way to one's knowledge. Over a reasonable period of time, with some concerted effort toward the systematic use of known conceptual frameworks, one's understanding about the relational context of a theory would be considerably improved. Thus, research on the relational context of theory in research would go on side by side with curriculum research per se.

A second relational context is that of previous knowledge and research. Although, in many ways, this is inseparable from theory, an area of unknown effect could be explored if curriculum research were conducted in terms of adequate communication of endeavors among curriculum researchers. It might be well to study the effects of adequate communication on curriculum research. One such position relevant to this problem may be Sterling's (1959) suppositions about the failure of many researchers and editors to publish null results. It would be interesting to know the number of poor ideas that have been retested because of the failure to communicate previous null results.

The third area of relational contexts in curriculum research to be considered is that of the empirical setting of the activity. A slight modification of Gage's (1963) continuum of research can be used to illustrate this context:

Basic science: indifferent content	Basic science: relevant content	Educationally oriented problems	Classroom experimentation	Field testing	Installation of program
Pure			*Simulated*		*Applied*
			RESEARCH		

In terms of the writers' definition, curriculum research would probably not be involved in the first two categories of the pure end of the continuum. As the continuum moves toward the applied pole, there is a marked change in the relational or situational context within which the research is done. Thus, one may move from studies of problems of learning in laboratories to the installation of a new mathematics program based on newly discovered principles of learning. Not at all clear, at this point, is whether a given aspect of an educational concern needs to progress through all stages, many stages, or a few stages. Of inestimable value would be knowledge of the effects of situational contexts on the findings of curriculum research.

A further related problem of changing contexts is the question of what should be done. Whether one likes it or not, curriculum research tends to go beyond the problem of what *is* and into the problem of what *should* be done. Curriculum decisions involve more than data; they involve values and other reality (e.g., practical problems). Thus, the application of a statistical test of significance to a problem in curriculum research does not end the problem-solving cycle. In other words, although the curriculum researcher organizes his research design on a what *is* basis, he cannot forget that his activity is applied research. He must remember that the cycle of problem solving will not be complete with a significant finding. There would appear to be an extremely valuable area of concern for curriculum research on the problem of how practitioners can move from results to applications in the most efficient and most useful manner. In other words, there should be some research concerned with the broad logic of applied research, which attempts to fit what *is* studies into what *should* contexts.

Problems of Design

As mentioned previously, most of the concern about designs used in curriculum research has been manifested on the verbal level rather than on the level of inquiry and research. For instance, Strang (1962) discussed weaknesses in the matched-groups approach as well as in correlational or factor-analytic studies. She suggested a single-group experiment in which all factors in a given set are described. Whether this is a feasible alternative or not, Strang did focus attention on the problem of design in curriculum research.

In this context, there have been many serious questions raised about the null hypothesis model for decision making, as evidenced in the papers by Green and Tukey (1960), Kaiser (1960), and Nunnally (1960). Underlying these questions is the conclusion that, since the null hypothesis is almost never true, an experiment to reject it is therefore unreasonable. Such arguments continue by asserting that, if an experiment fails to reject the null hypothesis, one conclusion, which is almost always possible, is that the sample was too small. Although there is a great deal of merit in the logical evidence attacking this model, research studies testing its efficiency as a decision-making model are lacking. Perhaps Sidman (1960) offers some direction to curriculum research workers when he espouses research focusing on individual behavior patterns. He asserts that group designs may fail to identify similar patterns among groups that have different means. Probably Sidman's most telling argument is his questioning of the process of randomization of uncontrolled variables.

Another pertinent area that involves necessary decisions by researchers is the establishment of levels of significance. More and more researchers are taking a hard look at the sacred cows of decision making—the .01 and .05 levels. Winer (1962), for one, suggested that, at times, even the .20 or .30 levels of significance might be appropriate. This area of concern could be one of legitimate interest to researchers. An attempt to establish some guidelines for setting up appropriate levels of significance in terms of sample sizes, powers of statistical tests, and other factors certainly seems advisable. This problem is also related to the context of communication mentioned earlier. Perhaps decisions about significance of findings should be left to the users of a study rather than to researchers or editors.

In an article directed at the central concern of this chapter, Lambert, Miller, and Wiley (1962) cited three requirements for an experimental design in curriculum research: (a) an adequate sampling of a research population, (b) introduction of treatment conditions that could be expected to maximize differences among treatment groups, and (c) use of sufficient numbers of subjects in each of the experimental groups. This last requirement is obviously related to the concept of power of statistical tests—a quantity that varies directly with the size of the experimental groups. Another pertinent question which, according to Lambert, is dealt with by all curriculum researchers either implicitly or explicitly is how much of an observed difference is required if it is to be judged significant. Of course, this question also directs the researcher's attention on the power of the statistical tests in his design.

The assertions with reference to curriculum research designs that have been stated in this section are, in the main, a result of conjecture. It is important that these assertions be tested in empirical settings that afford appropriate experimental designs. Of considerable aid to research workers in curriculum is Campbell and Stanley's (1963) chapter in the *Handbook of Research on Teaching*, in which 16 experimental designs are treated within the framework of 12 common threats to statistical inference.

Other Issues in Curriculum Research

Strang (1962) suggested that education is not yet a science and that premature attempts of educators to be scientific have led perhaps to misleading conclusions. Her argument seems to be focused on the rather weak or untested evidence-gathering techniques which researchers use.

In a vein similar to Strang's, Carroll (1961) identified the regressive maneuvers which educational researchers tend to fall back on—the use of questionnaires, supervisors' ratings, self-inventories, and the methodologies of statistics. He asserted that a direct attack is needed on the crucial variables of learning rates and attitudes, motivation, quality of instruction, application of the laws of learning, and teacher competence. Educators readily acknowledge that many factors affect behavior and that these many factors act in an interdependent fashion. It is interesting to note that educators rarely have definitive evidence to suggest which variables of concern are really independent and which are dependent. Certainly, specific variables may play either role in various research designs in terms of the labels assigned by researchers. However, the act of calling variables independent certainly does not assure that they are truly independent. In spite of this apparent insight into the complex relationships of variables, curriculum experiments generally concern themselves, as Fisher and Roth (1961) have noted, with single factors, single instruments, and single dimensions.

Curriculum researchers appear to be caught in a dilemma. Either they must wait for some theory or method that will allow them to deal with *all* crucial variables in curriculum research, or they may continue to work with a single factor or with several factors in anticipation of the ultimate and inevitable critical questions about uncontrolled or overlooked variables being raised at the close of the experiment. Perhaps Strang's (1962) and Sidman's (1960) views would lead one to suggest that, in curriculum research areas, experimental designs be waived temporarily. Curriculum researchers could resort to careful observations and descriptions until such time as the datum of curriculum research is identified and generally accepted by researchers. Then, meaningful hypotheses could be generated and tested.

The position regarding curriculum research, as defined in this chapter, may find some support in Gage's (1963) observation that use of measures of pupil gain or pupil growth as the ultimate criterion of teacher effectiveness is not feasible, at this time, in light of both conceptual and practical difficulties in use of the pupil gain criteria. Therefore, it may be reasonable to hold curriculum research at an observational stage until investigators concerned with research on teaching and learning find improved ways to understand and make predictions relevant to variables in these areas. Once this is done, it would seem that more fruitful research in curriculum could be undertaken.

Gage's position is quite relevant, in another sense, to researchers in the curriculum field. What is the ultimate criterion for judging the effectiveness of new teaching methods, content approaches, or administrative arrangements? The ultimate and ideal criterion would seem to be, as Gage suggests it is for teacher effectiveness, pupil gain or pupil growth. However, on closer examination, this issue seems to be highly complex. What of the findings of the television studies that reported essentially null results concerning achievement differences between groups of students in television and in nontelevision classes? Does this evidence suggest that one abandon television as an effective innovation, or, instead, that one should have all instruction via television? What other factors should be assessed? Since this criterion issue is of paramount importance, it needs to be explored by curriculum research workers. Unlike other issues raised in this chapter, empirical sources may not be adequate for validation of an assertion in this area. It may well be that a psychological orientation involving a systematic consideration of values will best serve as a basis for answering this question.

Conclusion

Although there is almost no research on curriculum research, the limited evidence presented from a review of the literature would seem to indicate that there certainly are a multitude of problem areas that could profitably be examined.

The area of relational contexts that involves relating reseach to theory, communicating knowledge from previous research, and perceiving the broad contexts of an applied discipline needs exploration and development. Problems of design cannot hope to be solved until comparative studies of different designs have been attempted. Indeed, the whole problem of experimental designs in contrast to careful observation seems appropriate for study. The need for common concepts, valid sources of evidence, and relationships of relevant variables, including the criterion variables, seems to be open to rigorous examination within any design problem.

Basic to all curriculum research is the question of whether research models taken from other fields are appropriate for curriculum research. Perhaps all, or at least some, are appropriate, and then again maybe there is a need for new curriculum research models per se. The only way in which these questions can be answered is through research on curriculum research. It seems self-evident that curriculum researchers should give this activity high priority.

Bibliography

Campbell, Donald T., and Stanley, Julian C. "Experimental and Quasi-Experimental Designs for Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. Chapter 5, pp. 171-246.

Carroll, John B. "Neglected Areas in Educational Research." *Phi Delta Kappan* 42: 339-43; May 1961.

Coladarci, Arthur P. "Toward More Rigorous Educational Research." *Harvard Educational Review* 31: 3-11; Winter 1960.

Fisher, Margaret B., and Roth, Robert M. "Structure: An Essential Framework for Research." *Personnel and Guidance Journal* 39: 639-44; April 1961.

Gage, N. L., editor. "Paradigms for Research on Teaching." *Handbook of Research on Teaching.* A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. Chapter 3, pp. 94-141.

Getzels, Jacob W., and Guba, Egon G. "Social Behavior and the Administrative Process." *School Review* 66: 423-41; December 1957.

Green, Bert F., Jr., and Tukey, John W. "Complex Analyses of Variance: General Problems." *Psychometrika* 25: 127-52; June 1960.

Herrick, Virgil E. "The Foundation Areas and Curriculum Development: A Summary and Interpretation." *Research Frontiers in the Study of Children's Learning.* (Edited by James B. MacDonald.) Milwaukee: School of Education, University of Wisconsin, 1961. pp. 36-46.

Joncich, Geraldine. "Whither Thou, Educational Scientist?" *Teachers College Record* 64: 1-12; October 1962.

Kaiser, Henry F. "Directional Statistical Decisions." *Psychological Review* 67: 160-67; May 1960.

Lambert, Philip; Miller, Donald M.; and Wiley, David E. "Experimental Folklore and Experimentation: The Study of Programmed Learning in the Wauwatosa Public Schools." *Journal of Educational Research* 55: 485-94; June 1962.

Nunnally, Jum. "The Place of Statistics in Psychology." *Educational and Psychological Measurement* 20: 641-50; Winter 1960.

Sidman, Murray. *Tactics of Scientific Research: Evaluating Experimental Data in Psychology.* New York: Basic Books, 1960. 428 pp.

Sterling, Theodore D. "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance or Vice Versa." *Journal of the American Statistical Association* 54: 30-34; March 1959.

Strang, Ruth. "Reactions to Research on Reading." *Educational Forum* 26: 187-92; January 1962.

Willower, Donald J. "Some Illustrated Comments on Hypothesis Construction and Research." *Journal of Educational Research* 56: 210-13; December 1962.

Winer, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill Book Co., 1962. 672 pp.

Additional References

Alexander, William M. "Assessing Curriculum Proposals." *Teachers College Record* 63: 286-93; January 1962.

Clymer, Theodore. "Research Design in the Language Arts." *Elementary English* 39: 349-54, 356; April 1962.

Evans, Herbert M. "Educational Research and Science Education." *School Science and Mathematics* 61: 601-606; November 1961.

Fleming, Robert S. "Improving Instruction Through Curriculum Research." *School Review* 66: 418-29; Winter 1958.

Heathers, Glen. "Field Research on Elementary School Organization and Instruction." *Journal of Educational Sociology* 34: 338-43; April 1961.

Jensen, Arthur R. "The Improvement of Educational Research." *Teachers College Record* 64: 20-27; October 1962.

KAZAMIAS, A. M. "Some Old and New Approaches to Methodology in Comparative Education." *Comparative Education Review* 5: 90-96; October 1961.
PELLA, MILTON O. "Criteria for Good Experimental Research in the Teaching of Science." *Science Education* 45: 396-99; December 1961.
SCHEFFLER, ISRAEL. "Justifying Curriculum Decisions." *School Review* 66: 461-72; Winter 1958.
STANLEY, JULIAN C. "Controlled Experimentation in the Classroom." *Journal of Experimental Education* 25: 195-201; March 1957.
WOODRUFF, A. C. "Educational Research and the Curriculum." *School Review* 66: 402-17; Winter 1958.

CHAPTER X

Curriculum Planning and Development: A Review of Reviews

WILLIAM M. ALEXANDER

IN THE first "Review of Reviews" to appear in the February 1962 REVIEW, Harris stated that such a chapter should "present a judgment of the extent to which the issue provides more than an abstracting service—valuable as that service may be" (p. 103). The present reviewer of reviews sees this judgment, for this particular issue, as answering five questions: What studies reviewed seem to be critically significant? What new and promising areas of curriculum research appear to have been opened? What contributions have been reported to a theory of curriculum planning and development? What new and promising procedures of curriculum research have been identified? What is ahead in research for curriculum planning and development?

In answer to the last question, in capsule form, the reviewer is definitely optimistic about the future of curriculum research. This area, notable mostly for neglect since the 1930's, is coming to life again. True, little within itself that is worthy of note has been concluded during the past three years. Indeed, this reviewer can say again what he did in introducing the June 1954 curriculum issue of the REVIEW: "the chief value of the present issue is to point up the need for a more significant effort on the part of educators to undergird their curriculum theories and practices with specific research studies" (p. 193). Now, nine years later, a statement can be made that was not made in the June 1954 issue: namely, significant beginnings have been made that are prophetic of meatier curriculum issues of the REVIEW in the future. By June 1969, if all goes well with cooperative research efforts now initiated or being planned, one would be privileged indeed to review the reviews of curriculum research. If the 15 years ending in June 1969 witness the rebirth and flowering of curriculum research, the profession will be able to take great pride in the accomplishments of these years.

Studies of Critical Significance

Probably the most promising research study under way in terms of its implications for curriculum planning and development is one not reviewed in the current issue. Project Talent, which was initially reported by Flanagan and others (1962), is in its infancy, of course, as far as publications are concerned. Perhaps because of this and, more likely, because of its great breadth, the project was not reviewed in any previous chapter of this issue. This reviewer, however, is thoroughly in agreement with Mouly (1963), who includes in his annotations of 21 selected significant research studies only one study initiated since World War II, namely, Project Talent. Although neither he nor the present reviewer would exclude all other studies of the past 15 years, certainly no project of recent or current times offers more hope for significant research than the great accumulation of data available through Project Talent.

It should be noted, too, that many other research studies supported by the Cooperative Research Program of the U.S. Office of Education, which are just under way, are not reported in the educational press yet. One can hope that many of these studies will offer significant data for curriculum planning and development.

Authors of the various chapters identified a few studies that offer findings or hopes of findings of real significance. For example, the authors of Chapter V observed that there is much research dealing with the work of the School Mathematics Study Group (SMSG), although the research currently concluded has barely scratched the surface. On the other hand, it is gratifying to see the scope of the research already conducted and to note the report that SMSG and the Minnesota National Laboratory have planned a new and comprehensive five-year study. If these new studies can eliminate the deficiencies of those already concluded, then we may be hopeful about the prospects of significant research related to nationally oriented curriculum studies.

The authors of Chapter III reviewed many studies that have interesting implications for curriculum planning. Unfortunately, as they observed, some of these studies simply support or replicate older research. On the other hand, this reviewer finds some optimism about the research reported, particularly with reference to socioeconomic class. Such studies as those by Coleman (1961) and Havighurst and others (1962) have great potential for the modification of curriculum planning and development. As a matter of fact, one wonders whether research in such areas as the relationship of socioeconomic class and education and the relationship of the peer culture to social and educational aspirations might not need replication frequently in differing communities—perhaps periodically in every community in order to provide basic data for designing educational programs.

Although some might question the inclusion of a chapter on teaching in a curriculum issue of the REVIEW, the authors of this chapter in the present issue identified some research believed by the reviewer to have extreme significance for the planning and development of classroom experiences, a major element of the curriculum. If the quality of pupil-teacher interaction is a critical factor in the curriculum of the individual, then what can be determined about the teacher's role in this interaction is important in curriculum planning and development. Hence, such publications reviewed in Chapter VI as those by Barr (1961), Flanders (1961), Gage (1963), Hughes (1962), Ryans (1960), and Smith (1961) affirm the optimism with which one views the present status of curriculum research.

Perhaps, however, the greatest basis for optimism concerning the field is to be found in the studies reviewed in Chapter VIII. Interestingly enough, the authors of Chapter VIII seem to view the development of models with more satisfaction than do the authors of Chapter I. Could it be that more has been learned about change in the process of education than about the nature of the process itself? References to studies cited in Chapter VIII are made at further points in this chapter.

New and Promising Areas of Curriculum Research

Although not new, perhaps the most significant development in curriculum planning in recent years is the re-examination of the disciplines as sources of curriculum content. Although the author of Chapter II was unable to identify significant research in this area, he did review promising efforts to examine the structure of the disciplines and possible approaches to their utilization in curriculum planning. He pointed out that "present results may be viewed, for purposes of research, as an inventory of propositions to be investigated." The analysis of the disciplines is a rich and promising area for research that can weld together once again the scholarship of the university and that of the elementary and secondary schools.

The studies reviewed in Chapter III give a new emphasis to the need for research to identify curriculum provisions that may raise the social status and aspirations of many American youths. Although existing research findings offer too little guidance for elementary and secondary schools in meeting such problems as the dropout and the juvenile delinquent, the profession seems to know more about the backgrounds of these problems than ever before and to exhibit more interest in dealing with them intelligently. Thus, one may well hope that continued and greatly increased attention will be given to the remediation and prevention of such social ills through deliberately conceived and experimentally tested educational programs.

The authors of Chapter VIII reported a number of leads to new and promising research areas. For example, the strategy of curriculum change proposed by Taba (1962) has never been carefully tried out on an experimental basis. This *inversion approach* (the present reviewer prefers the term *inductive approach*) to beginning curriculum planning and development experimentally through specific teaching-learning units could well be the subject of relatively simple and extensive research. The authors of Chapter VIII also offer a provocative suggestion about a new and promising area by proposing that "educators might well explore the impact of the structure of the profession itself upon curriculum development and curriculum change." Although such research might be highly complex, information that could be secured concerning the relationship of the professional power structure to the educational program itself might lead to significant deviations from the traditional alignment of professional personnel and controls.

New and Promising Research Procedures

It is, of course, very difficult to judge research procedures on the basis of reading only the reviews presented in the preceding chapters of this issue. On this basis only, the reviewer is of the opinion that the only new and promising research procedures clearly identifiable are those with reference to the area of Chapter VI, "Teaching." Great ingenuity was exhibited in a number of the research studies reported in this chapter. The interaction analysis used by Flanders (1961) to determine the degree and effects of teacher influence and items of teacher and student behavior seems of utmost significance in continued study of the teacher-pupil interaction element of the curriculum. Although the technique of analysis of verbatim records of classroom conversation is not new, the employment of this procedure by Hughes (1962) and by Smith (1961) as a basis for description and classification of verbal behavior in the classroom seems to offer a procedure for continuing analysis and evaluation of classroom experiences of pupils. The work of Ryans (1960) in using various techniques for the study and evaluation of teacher characteristics in many situations is suggestive of a team research approach to studies of specific school activities in a variety of situations.

The use by Dalton (1962) of pupil ratings to differentiate effective and ineffective teachers and to analyze differences in the preparation of these teachers suggests a procedure of possible significance in the study of many aspects of teacher backgrounds and performance. The more extensive use of relatively new instruments in the analysis of classroom teaching and pupil-teacher interaction should also be noted.

The authors of Chapter VIII made a very interesting suggestion about the dissemination of educational research: "some type of dissemination service or agency that would function in a manner comparable to the county agent in agriculture to help teachers and administrators relate the best research results available to the local school situation." Perhaps, a true breakthrough in education could be greatly aided by establishing, in a large number of school systems, the position of research communicator.

The reviewer would like to call attention to the interesting and careful method used by the authors of Chapter V in the preparation of their material. Application of standards used in this chapter, such as selection of items for review as well as the check by telephone conversations with persons actively engaged in major research projects, would probably improve the quality of many chapters in future issues of the REVIEW.

A Theory of Curriculum Planning and Development

The authors of Chapter I found great need for a theory of curriculum planning and development. They identified some efforts in this direction but, as yet, no significant formulation of a comprehensive theory. They concluded that the job is yet to be done and indicated stages and steps suggested by the publications reviewed.

In the reviewer's judgment, the authors of Chapter VIII noted the greatest current contribution to a theory of curriculum planning and development in their treatment of the "diffusion of innovations." They demonstrated their critical approach to this matter through clear-cut criticisms of the limitations of the Brickell (1961) report. At the same time, they recognized the broader view of diffusion research taken by Rogers (1962) and also the comments of Mackenzie and Huebner (1962) on the political qualities of curriculum change. Such studies as these indicate the tremendous possibilities of more rapid and comprehensive utilization of significant research on the curriculum once the latter is available.

The authors of Chapter VIII also presented reviews of some of the significant work done in the studies of roles, role perception, and personality structure. Although one might expect to find some of these studies as well as those of creativity reviewed in Chapter IV (which the present reviewer did not have opportunity to read in the preparation of Chapter X), nevertheless, such work as that of Flanders (1960) and of Bills (1963) offers significant insights into the operation of curriculum planning organizations. Certainly, the ways in which the individuals concerned view themselves and their colleagues must occupy a central place in any theory of curriculum planning and development.

Looking Ahead

As indicated in the introduction to this chapter, the reviewer is well aware of the paucity of significant, completed curriculum research studies. He quite agrees with the authors of Chapter V that "too few studies are replicated" and that "too few studies have an end beyond themselves." He is impressed, too, with the final recommendation of these authors: "This journal could perform a vital service by rigorously selecting fewer studies for review, by analyzing each study in depth, and by candidly reporting its strengths and weaknesses."

This final chapter, however, has attempted to indicate the reviewer's judgment of some of the promising beginnings of significant research and the development of critical areas and ideas for exploration. His optimism is most aroused by a seeming ground swell of interest in, even beginnings of, long-term, carefully conducted, curriculum-oriented research studies.

Bibliography

Barr, A. S., issue editor. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations." *Journal of Experimental Education* 30: 5-156; September 1961.

Bills, Robert E. "Education in Human Relations." *New Insights and the Curriculum.* 1963 Yearbook. Washington, D.C.: Association for Supervision and Curriculum Development, a department of the National Education Association, 1963. Chapter 8, pp. 165-87.

Brickell, Henry M. *Organizing New York State for Educational Change.* Albany: New York State Department of Education, 1961. 107 pp.

Coleman, James S. *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education.* New York: Free Press of Glencoe, 1961. 368 pp.

Dalton, Elizabeth L. *What Makes Effective Teachers for Young Adolescents?* Nashville: Department of Education, George Peabody College for Teachers, 1962. 34 pp.

Flanagan, John C., and others. *Design for a Study of American Youth.* Boston: Houghton Mifflin Co., 1962. 240 pp.

Flanders, Ned A. "Analyzing Teacher Behavior as Part of the Teaching-Learning Process." *Educational Leadership* 19: 173-75, 178-80; December 1961.

Gage, N. L., editor. *Handbook of Research on Teaching.* A project of the American Educational Research Association, a department of the National Education Association. Chicago: Rand McNally & Co., 1963. 1218 pp.

Havighurst, Robert J., and others. *Growing Up in River City.* New York: John Wiley & Sons, 1962. 189 pp.

Hughes, Marie M. "What Is Teaching? One Viewpoint." *Educational Leadership* 19: 251-59; January 1962.

Mackenzie, Gordon N., and Huebner, Dwayne. "Politics of Curriculum Change." *Curriculum Crossroads.* (Edited by A. Harry Passow.) New York: Bureau of Publications, Teachers College, Columbia University, 1962. pp. 76-86.

Mouly, George J. *The Science of Educational Research.* New York: American Book Co., 1963. 515 pp.

Rogers, Everett M. *Diffusion of Innovations.* New York: Free Press of Glencoe, 1962. 367 pp.

Ryans, David G. *Characteristics of Teachers.* Washington, D.C.: American Council on Education, 1960. 416 pp.

Smith, B. Othanel. "A Concept of Teaching." *Language and Concepts of Education.* (Edited by B. Othanel Smith and Robert H. Ennis.) Chicago: Rand McNally & Co., 1961. Chapter 6, pp. 86-101.

Taba, Hilda. *Curriculum Development: Theory and Practice.* New York: Harcourt, Brace and World, 1962. 529 pp.

Index to Volume XXXIII, No. 3

Page citations are made to single pages; these are often the beginning of a chapter, section, or running discussion dealing with a topic.

Analyses and implications of disciplines: English, 248; mathematics, 248; science, 248; social sciences, 248
Automated instruction and teaching machines: 271

Chart or charter, recent developments in educational discourse: 233
Chemistry: 282
Criticisms of educational research: 233
Curriculum and communications media: merchandising and research, 302; technology of teaching, 301
Curriculum components and organization: curriculum organization, 288; English, 284; foreign languages, 287; mathematics, 278; social studies, 287
Curriculum development, dynamics of change in: contributions from related fields, 315; diffusion of innovations, 306; field demonstration, 308; need for further research, 316; redefinition of change, 310; roles, role perception, and personality structure in change, 311; search for models, 304
Curriculum organization: team teaching, 288; track organization, 289
Curriculum planning and development: new and promising areas of curriculum research, 332; new and promising research procedures, 333; studies of critical significance in, 330; theory of, 334
Curriculum research: issues in, 326; problem of design, 324; problem of relational contexts, 322

Disciplines: analyses and implications of, 247; curriculum source recognized, 245

Educational discourse, recent developments in: 233
Educational research, criticisms of: 233
English: review of studies, 285; theory of language instruction, 284

Forces influencing curriculum decisions: organized knowledge, 245; psychological, 268; societal, 254
Foreign languages: 287

Group interaction: 271

Mathematics: Educational Testing Service study, 279; Minnesota National Laboratory studies, 279; SMSG research, 280
Models: in general, 240; search for, 304

New and promising areas: in curriculum research, 332; in research procedures, 333

Personality structure in relation to dynamics of change in curriculum: 311
Physics: 281
Problems of design in curriculum research: 324
Psychological forces influencing curriculum decisions: learner, 268; learning outcomes, 272; learning process, 269; research design and psychological influences, 273; school learning situation, 270

Relational contexts in curriculum research: 322
Research designs and psychological influences: 273
Research procedures, new and promising: 333
Research theory, nature of: analysis of problems, 235; natural history stage, 236; stages of inquiry, 234
Roles and role perception in relation to dynamics of change in curriculum: 311

School learning situation: automated instruction and teaching machines, 271; group interaction, 272; task and method variables, 271; teacher, 270
Science: chemistry, 282; physics, 281
Social problems in relation to public schools: delinquency, 260; dropouts, 259; economic influences, 262; political problems, 261
Social realities in relation to school situation: family life, 258; peer culture,

257; racial and ethnic differences, 254; religion, 255; socioeconomic class, 256
Social studies: 287
Structures and groupings of disciplines: 249
Studies of critical significance in curriculum planning and development: 330

Task and method variables: 271
Teacher: 270
Teacher and teaching effectiveness: characteristics, 296; personality, 295
Teaching: analysis of teaching act, 293; teacher and teaching effectiveness, 295; technology of, 301
Team teaching: 288
Technology of teaching: 301
Theory for education: defined, 239; in educational research, 238; in relation to curriculum development, 238; usages of the term, 237
Theory of curriculum planning and development: 334
Theory making in education: incomplete analogy, 242; models, 240
Track organization: 289

REVIEW OF EDUCATIONAL RESEARCH

Official Publication of the American Educational Research Association
Contents are listed in the Education Index.

Copyright 1963
By National Education Association of the United States, Washington, D.C.

Volume XXXIII, No. 4 **October 1963**

Teacher Personnel

Reviews the literature for the five-year period since the issuance of Vol. XXVIII, No. 3, June 1958.

TABLE OF CONTENTS

Chapter *Page*

Foreword 342

I. Teacher Certification, Supply, and Demand 343
JAMES C. STONE, *University of California, Berkeley, California*

II. Recruitment and Selection of Prospective Elementary and Secondary School Teachers 355
GLENN W. DURFLINGER, *University of California, Santa Barbara, California*

III. Pre-Service and In-Service Education of Teachers 369
HAROLD E. REYNARD, *Ohio State University, Columbus, Ohio*

IV. Conditions of Employment and Service in Elementary and Secondary Schools 381
GARFORD G. GORDON, *California Teachers Association, Burlingame, California*

V. Teacher Personnel Research in Higher Education 391
ALVIN MARKS, *Sonoma State College, Cotati, California*
MARVIN N. DILLON, *Sonoma State College, Cotati, California*

Chapter *Page*

VI. Economic, Legal, and Social Status of Teachers 398

Hazel Davis, *National Education Association, Washington, D.C.*
Martha L. Ware, *National Education Association, Washington, D.C.*
Frieda S. Shapiro, *National Education Association, Washington, D.C.*
Eleanor Donald, *National Education Association, Washington, D.C.*
Gertrude N. Stieber, *National Education Association, Washington, D.C.*

VII. Assessment of Teacher Behavior and Instruction 415

David G. Ryans, *System Development Corporation, Santa Monica, California*

VIII. Teacher Personnel: A Brief Evaluation of the Research Reviewed 442

William B. Michael, *University of California, Santa Barbara, California*

Index 449

This issue of the REVIEW was prepared by the
Committee on Teacher Personnel

HOWARD A. BOWMAN, *Chairman,* Los Angeles City Schools, Los Angeles, California
HAZEL DAVIS, National Education Association, Washington, D.C.
GLENN W. DURFLINGER, University of California, Santa Barbara, California
GARFORD G. GORDON, California Teachers Association, Burlingame, California
ALVIN MARKS, Sonoma State College, Cotati, California
WILLIAM B. MICHAEL, University of California, Santa Barbara, California
HAROLD E. REYNARD, Ohio State University, Columbus, Ohio
DAVID G. RYANS, System Development Corporation, Santa Monica, California
JAMES C. STONE, University of California, Berkeley, California

with the assistance of

MARVIN N. DILLON, Sonoma State College, Cotati, California
ELEANOR DONALD, National Education Association, Washington, D.C.
FRIEDA S. SHAPIRO, National Education Association, Washington, D.C.
GERTRUDE N. STIEBER, National Education Association, Washington, D.C.
MARTHA L. WARE, National Education Association, Washington, D.C.

FOREWORD

Up to June 1958, research in teacher personnel had been reviewed at three-year intervals. The present issue of the REVIEW is concerned with research in the same topic over a five-year period.

With respect to the studies of teacher personnel, chairmen of the June 1955 and June 1958 issues of the REVIEW agreed that there was little good research to be reviewed and that there was need for further study and for development of new techniques. To a large extent, this is still true.

Some portions of the subject studied did not lend themselves to research other than the descriptive or normative types. Other portions, however, proved to be sources of both intensive and extensive research during the past five years. In two or three instances, the amount and quality of research was more than could be reviewed adequately, due to limitations of space and time. For this reason, a return to the three-year cycle is suggested.

The next committee to review research on this topic may wish to undertake some internal reorganization of the subject. From what has been discovered by the present chapter authors, it seems likely that the contents of each of two of the chapters might well have been organized as four separate chapters. On the other hand, certain other chapters might well have been combined into one. Unfortunately, the quality and volume of research did not reveal themselves until work was well under way, at which time it was too late to attempt major reorganization.

HOWARD A. BOWMAN, *Chairman,*
Committee on Teacher Personnel

CHAPTER I

Teacher Certification, Supply, and Demand

JAMES C. STONE

TEACHER certification, supply, and demand are so interrelated that separation for more than descriptive purposes tends to neglect the most important aspects of analysis. The pervading theme throughout the literature reviewed attests to the need for an adequate supply of fully qualified personnel in the public schools. The issue is compounded, however, for there is no precise agreement on the meanings of any of the terms cited, although the National Commission on Teacher Education and Professional Standards (NCTEPS) and other units of the National Education Association have been working for common practices and understandings.

Certification

Requirements for regular certification in some states actually have been lower than those for emergency certification in others, according to Hodenfield and Stinnett (1961). This has been a major problem of the whole situation of certification in the United States. If the inequalities between emergency and standard certificates emerge as a problem, that of the comparability of standard certificates among states is one of truly significant proportions.

State-by-State Variations

Armstrong and Stinnett (1962), Hodenfield and Stinnett (1961), and Woellner and Wood (1962) reported specific differences in standard certificates for elementary and secondary teachers among different states. Arizona, California, and the District of Columbia require a master's degree or 30 semester hours of postgraduate work for a high school certificate; Connecticut, Indiana, Maryland, New York, Oregon, Pennsylvania, and Washington require a fifth year of preparation for full certification for secondary school teachers (Armstrong and Stinnett, 1962). Arkansas is the only state not requiring a bachelor's degree for secondary teaching, although it is planned as a state requirement sometime in 1963. The addition of Arkansas in 1963 will represent a three-state gain since Herge (1958) completed his review of certification requirements.

Woellner and Wood (1962) reported that in only six states the bachelor's degree is not a minimal requirement for regular certification of elementary school teachers. Since 1958, seven states—Alaska, Colorado, Iowa, Kansas, Kentucky, Minnesota, and Missouri—have changed to a minimum standard of a bachelor's degree for elementary teachers. Those states not requiring this minimum—Maine, Montana, Nebraska, North Dakota, South Dakota, and Wisconsin—tend to be somewhat sparsely populated with smaller and more isolated school districts (Armstrong and Stinnett, 1962).

The trend toward the requirement of a bachelor's degree also was recognized within the professional organizations of at least eight states by the end of 1961 (Armstrong and Stinnett, 1962). One anticipates that this trend will continue at the state level as more states develop higher standards for full certification.

Variations in Certificates

Within the framework of eligibility for standard certification, other major points of divergence among states pertain to the nature of the specific requirements in professional education, subject matter, specialization, general education, citizenship, special courses, and loyalty oaths. In discussing several of the inequities among states, Barnes and Shipman (1961) found that the requirements in professional education for elementary school teachers in 1960 ranged from 16 semester hours in three states to 36 semester hours in one state and that for secondary teachers the same kind of requirements varied from 12 semester hours in two states to 27 semester hours in one state. Moreover, the number of hours required for student teaching fell between zero and 10 semester hours.

Barnes and Shipman also concluded that general education requirements, which are not specified in most states, continue to receive less attention than areas of subject matter specialization. They also noted that requirements in the specialized subject matter of the applied arts and sciences were higher in number than were the subject matter requirements in the traditional academic areas in high school. In a corresponding study, Tuttle (1963) determined that the applied fields were requiring two to three times as much preparation as the traditional fields. He urged that the differences in requirements be narrowed. Of a related nature was Armstrong and Stinnett's (1962) report that the most significant recent change in certification has been the general upgrading by the states of the subject matter requirements in the traditional academic areas.

Levels of Responsibility

State boards of education are the most common agencies for the issuance of certificates, according to Armstrong and Stinnett (1962), Barnes and Shipman (1961), Latimer (1963), and Woellner and Wood (1962). Although Barrett (1961) found that centralization of authority for certification in the New Mexico State Board of Education was associated with the higher standards increasingly being demanded of the state's teaching personnel. Lindsey (1961) supported the thesis that the profession itself must take more responsibility for the enforcement and administration of the licensing system than it has taken in the past. Evidence that this point of view is gaining momentum may be noted in the writings of Allen (1961), Barnes and Shipman (1961), Bush (1961), Corey (1961), Davies (1962), Latimer (1963), and Yeager (1960).

One of the most pronounced indications of the increasing amount of responsibility being placed upon the profession has been through the approved program approach for meeting certification requirements. It is through this approach that NCTEPS (Lindsey, 1961) has attempted to establish guidelines and standards which provide an increasing degree of reciprocity among the states. Barnes and Shipman (1961) reported a Northeast 11-state compact for reciprocity and a Midwestern agreement that provides freedom of movement across state boundaries with minimal difficulty. Davies (1962) mentioned that 27 states have agreed to accept for certification any graduate of an institution approved by the National Council for Accreditation of Teacher Education (NCATE). Lander (1963) noted that approximately 400 of the 1,100 colleges and universities offering teacher education programs have the endorsement of NCATE and that more than one-half of these institutions have entered into reciprocal agreements for certification of teachers on the approved program basis. Latimer (1963) declared that complete reciprocity in certification of teachers is present in 20 states. Authors who reported upon approved programs were quick to point to the added responsibility that such a process placed on the institutions for careful selection and screening for teaching credentials.

Teacher Examinations

A number of writers discussed various certification approaches that utilize some form of examination of teaching competency. Davies (1962) called for examinations independent of the preparing institution and of the employing school district. Latimer (1963) reported that teacher certification from subject field to subject field should depend on scores on an examination that reflects the cooperative planning of professional organizations and learned societies. He also would restrict courses from being offered by local schools unless a qualified teacher was available to do the teaching. After calling attention to the West Virginia plan of using the *National Teacher Examinations* to test knowledge of the teaching field, Stiles (1959) reported a consideration by the Wisconsin State Department of Public Instruction to use written examinations for potential teachers. These examinations would furnish measures concerning breadth of liberal arts background, knowledge in specific subject fields, and familiarity with objectives in professional education courses. Taking a position different from that of Davies (1962), Stiles urged that the institution in teacher education maintain the responsibility for administering the examination and for interpreting the results—an activity that would furnish one basis of recommendation for certification. Yeager (1960) suggested an "extra-official, professional board," which would be closely allied, but not tied, to the state certification agency as the proper examining and recommending agency. The board would represent the profession in certification procedures. Greene (1961) also called for close cooperation of the profession and the state.

In noting that 22 states made use of some sort of examination for meeting various state certification requirements, Slaughter (1961) believed that the apparent trend in recent years toward the use of examinations must be the result of an attempt by the states to make certification procedures more flexible than they have been in the past.

Along with increased use of examinations has come a decrease in the number of states requiring special courses on a state or regional basis. However, Armstrong and Stinnett (1962) pointed out that there are still a number of states which require U.S. citizenship, a loyalty oath, a general health certificate, a chest X ray, or special courses for certification requirements. Woellner and Wood (1962) reported that there were 579 separate-name certificates issued by the states in 1961 and that the number per state ranged from 2 in three states to 68 in one state. This represents a decrease in the total number from 630 in 1959, as noted by Hodenfield and Stinnett (1961). Hodenfield and Stinnett also stated that the most common certificate issued is one that carries an endorsement indicating the field or area of specialty.

Trends

Many authors attempted to spot trends in certification of teachers. Although more states than formerly are accepting NCATE-approved programs as a basis for certification, the United States is still a long way from national certification (Hodenfield and Stinnett, 1961). Lieberman (1961) outlined the advantages of national certification and concluded that basic changes in the approach to certification would be necessary before such a procedure was possible. Voicing a view of NCTEPS, Lindsey (1961) reinforced the program plan approved by NCATE as a necessary aspect of professionalization and as a standard of certification on a national level.

Armstrong and Stinnett (1962) reported that all states require certification of all public school personnel who have specific responsibilities related to pupils. Included are administrators, nurses, guidance counselors, psychologists, and social workers. Within the general requirement that all personnel be certified, the National Education Association, Research Division (1962a) estimated that 91,556 teachers were employed on the basis of emergency certificates in the 1962-63 school year. This figure indicated that 1 teacher in 16 was serving on a substandard certificate. It was also estimated that approximately 70 percent of these emergency-certified teachers were employed in the elementary schools and 42 percent in rural schools. Approximately 47 percent did not have at least four years of college preparation. This fact takes on significance when one considers that six states do not require a bachelor's degree for full certification for elementary school teachers. The National Education Association, Research Division (1963b) reported that 6.3 percent of the elementary and 4.3 percent of the high school teachers had less than full teaching qualifications for their states.

The supply of qualified teachers for public schools in the United States remains the foremost problem in public education today. The National Education Association, Research Division (1963a) reported a total of 1,509,281 classroom teachers in the public schools for the 1962-63 school year. The Division also estimated a net shortage of 134,000 teachers for this same period.

Supply and Demand

Maul (1960b) called attention to the factors both within and without the educational system that have had an effect on the supply of and demand for qualified teachers. To give examples of several external forces, he pointed to increased birth rate, mobility of the population, concentration of young families in such temporary locations as areas containing plants or military installations, and desires of local communities to provide education over and above the state minimum. Maul also listed the internal conditions of high standards for teaching, which tend to contribute to stability and minimum turnover, as opposed to the low retention rate of communities with relatively low standards. He went on to generalize that economic factors, rather than just the number of qualified persons in the general population who are capable of teaching, constitute the most influential condition underlying the supply-demand problem.

Economic Factors

The importance of salary and of other economic considerations to the supply of teachers was substantiated in a large number of studies. Fox (1961), Harris (1957), Hills and Downs (1962), Lambert (1963), Maul (1961), McKean and Kershaw (1962), the National Education Association, Research Division (1962b), Richards (1960), Seltzer (1962), Stone (1961), Tyler and Miller (1962), Van Pelt (1958), and Walker (1958) provided data that supported Maul's emphasis on economic factors as important determinants of the supply-demand ratio.

Subject and Grade Level Factors

The demand for teachers is one not only of numbers but also of distribution of those who prepare for teaching. Maul (1960a, b, c, 1961), McKean and Kershaw (1962), the National Education Association, Research Division (1960a, b, 1962b, c), Osborn and Maul (1961), and Stone (1961) referred to the need for an improved balance between subject matter fields in greatest demand and those consistently oversupplied.

Bartels (1960) explained the Index of Teacher Demand for 1958 as the ratio of the number of new teachers employed who had not taught during the previous year to the number of teachers newly prepared by colleges and universities. It is expressed as the percentage which demand is to supply. Although Bartels presented information on only 29 states and 2 territories, his findings proved interesting. The Index for the high school was 83 percent; for the elementary school, 188 percent.

The National Education Association, Research Division (1960a, b, 1962b, c) has reported a continuing condition of eight new elementary positions available for every five high school positions; the ratio of supply has been the reverse. An encouraging trend in the differential between the numbers of elementary and secondary school candidates who prepared for teaching between 1950 and 1955 revealed an uncertain reversal during the years between 1955 and 1960. The National Education Association, Research Division (1962b, c) reported an increment of 13.9 percent of the new prospective high school teachers as opposed to an increase of 5.1 percent for new prospective elementary teachers over the previous year. This circumstance did not decrease the need for teachers in the elementary schools. Although only 68 percent of the high school and 83 percent of the elementary school trainees who prepared for teaching each year actually entered teaching, the difficulties of supply have become increasingly acute in certain fields of high school teaching.

A gross example of oversupply was in the area of men's physical education. Of the 7,700 graduates with major preparation in this field, fewer than one-quarter obtained full assignments in physical education; of the total graduates in physical education who obtained positions, more than 17 percent were teaching outside of their major area of preparation. Social science was next in the order in which supply exceeded demand. Homemaking education and women's physical education were at the other end of the scale, with demand far exceeding supply. The physical sciences and mathematics were also fields of continuing extreme shortage. These supply-demand ratios at the high school level were almost lost in the tremendous need for elementary school teachers. There appears to be little doubt that the area of greatest need will continue to be in the elementary field.

Other Factors

In addition to the economic factors, there have been other conditions that contribute heavily to the deficit in supply. The continuing increase of children of elementary and high school age shows few signs of abating. The National Education Association, Research Division (1963a), which reported 20 percent of the U.S. population to be in the school age bracket (between 5 and 17 years), estimated that by 1970 the population in this age group will be at least 6 million more than it was in 1950. It also was estimated that the same school age group will constitute 25 percent of the general population. Clearly, the need for teachers will increase.

Folsom (1958), Harris (1957), and Walker (1958) called attention to the indifferent or negative attitudes of some of the public toward teachers and emphasized the detrimental effect such attitudes have on keeping teachers in the classroom. Fox (1961), Maul (1960a), the National Education Association, Research Division (1960a, b, 1962b, c), Osborn and Maul (1961), Stone (1961), and Van Pelt (1958) called for efforts in counseling youth about the specific areas of need in teaching. Hills and Downs (1962) pointed out that the profession has failed to offer new teachers opportunities for advancement that are commensurate with their ability level. This shortcoming well may discourage many college students from entering teaching.

Maul (1960b) reported that a high standard not only reduces teacher turnover but also exerts a stabilizing effect on the educational system. In describing factors that contributed to teachers' withdrawing from Ohio and Oklahoma schools, respectively, Harris (1957) and Walker (1958) also found that high standards, good salaries, and good working conditions tended to go together.

The differences in state certification requirements, in salaries, and in working conditions also have an effect on state planning of supply and demand of teachers. The National Education Association, Research Division (1962c) noted that 15.7 percent of the graduates during the 1960-61 school year who were certified for teaching took jobs outside of the state in which they had prepared. This figure, compared to 14.9 percent, cited two years previously by the National Education Association, Research Division (1960b), accounted for one out of eight of the teachers in the group. This report also drew attention to the fact that states which had a high percentage of college graduates among the total elementary school corps of teachers were able to make further progress in recruiting highly qualified teachers, who, in turn, would tend to raise the total state average. On the other hand, those states with total elementary staffs at the lower levels of preparation tended to be inducting new teachers with less than the national average amount of preparation. This underscores an important trend—the widening of the gap between states with higher standards and those with lower standards. Coupled with this trend is the probability that the quality teacher in a low standard state can find greater economic rewards and better teaching conditions outside the state in which he was educated.

Both the National Education Association, Research Division (1962c) and Stone (1961) reported that approximately one-half of all persons certified each year in California have been prepared in out-of-state institutions. Stone (1961) also reported that California colleges and universities that were accredited to prepare teachers supplied only 52 percent of the total number of new teachers needed in 1960.

The mobility of teachers is a prime factor in the inability of researchers to assemble precise data on the supply and demand of teachers. Hodenfield and Stinnett (1961) reported on the difficulties teacher mobility brings to standardization in certification—a mobility which challenges the state agencies responsible for yearly estimates of supply and demand. The general movement of the population to the West and the Southwest and from nonmetropolitan to metropolitan areas, especially the suburbs, was reported by the National Education Association, Research Division (1961). Tyler and Miller (1962) called attention to the special need for teachers in growing urban areas.

Stone (1961) reported that one-third of the teachers in California had been educated outside the state. One-third had taught in other states, and only one-half had taught in only one California school district. He noted that a slightly higher percentage of high school teachers than of elementary teachers remained in one district. Somewhat disconcerting was the finding that 27 percent of the candidates trained in and certified by California failed to enter teaching.

Osborn and Maul (1961) noted that approximately 64 percent of the elementary school teachers prepared in Iowa remained to teach in that state and that approximately 28 percent went into elementary schools in other states. The remaining new teachers did not enter teaching. Approximately 27 percent of the candidates who prepared for the high school level did not enter teaching, and only 53 percent of those trained remained in Iowa.

Remedies for the Shortage

Several remedies for the shortage of teachers have been suggested. Folsom (1958) suggested retaining more of the qualified teachers already prepared. He stated that, if the profession could extend the average teaching career by two years, the shortage practically would be eliminated. With marriage and maternity reported by Harris (1957) and by Stone (1961) to be the two major causes for teachers leaving the profession, there is little immediate hope for Folsom's suggestion answering the need.

As implied previously, substantial increases in compensation would be the major step toward solving the shortage. McKean and Kershaw (1962) suggested that, since the shortages in teaching are specific, they should be approached through special salary incentives which would be offered in those fields of greatest shortage. They challenged the across-the-board salary schedule as highly expensive and, in fact, declared it a contributor to the continued disproportionate supply of teachers.

Such studies as those by Fox (1961), Richards (1960), Tink (1960), and Van Pelt (1958) attested to the importance of teacher influence on college students who entered teacher education programs. The tendency was for the elementary school teachers of these college students to have had the most marked influence.

Seltzer (1962) indicated that the high ratio of high school to elementary school teachers might be influenced by providing a subject matter major in the elementary school teacher education program. This was based on his finding that the interest in a particular subject field was an important factor in the choice of prospective teachers who entered secondary education programs.

Phi Delta Kappan (1963, p. 355) reported that in the next decade 50,000 men and women will retire annually from military service. Nearly one-half of these persons have bachelor's degrees. Moreover, two-thirds of the retiring officers have expressed an interest in education as a second career. Shriver (1963) stated that almost two-thirds of the 5,000 Peace Corps volunteers are working in education fields. He felt that the eagerness of these people to enter teaching might be diminished by certification requirements. He also indicated that various universities have plans to furnish Peace Corps members with educational and certification credit. *Phi Delta Kappan* (1963, p. 355), which reported a poll of the first 700 returning Peace Corps members, revealed that approximately 25 percent looked forward to teaching careers.

The solution to the quantity problem of supply and demand appears in sight for some states. Stone (1961) projected the California supply of teachers to surpass demand by the academic year 1970-71. This look to the future does not, however, take into account the need for replacing teachers who currently are serving on substandard credentials. Quality teachers appear to be in demand for some years in the future. *Phi Delta Kappan* (1963, p. 353) reported an interpretation by Coulter of national figures, in which he also visualized an easing of the teacher shortage by 1970. The interpretation was based on the fact that between 1960 and 1970 an increase in enrollment of 24 percent and an increase of 79 percent in the number of college graduates will take place. Coulter contended that the college graduate is the primary source of supply for the teacher shortage and that the increased number of graduates alone will be the chief factor in reducing the shortage.

Needed Research

Anderson and Rusher (1960) called attention to the lack of research on the value of certification practices. Other areas of needed research include the relationship of certification to the quality of teacher competence, the instrumental value of certification in the control of teacher education, and the sociological implications of national standards of competence and certification. The problems of supply and demand leave many questions unanswered. Research is needed on how to find out the information which is required for making accurate prognoses concerning fields of teacher demand and means of meeting them. The mobility both of the nation's population and of teachers suggests the need for improved methods of locating teachers and of keeping in touch with them over a period of years.

Bibliography

ALLEN, WENDELL C. "Goals of Certification." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 277-82.

ANDERSON, EARL W., and RUSHER, ELFREDA M. "Staff-Certification." *Encyclopedia of Educational Research.* Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1354-57.

ARMSTRONG, W. EARL, and STINNETT, T. M. *A Manual on Certification Requirements for School Personnel in the United States.* 1961 edition. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1962. 219 pp.

BARNES, DONALD L., and SHIPMAN, CHARLES D. "Changing Patterns in State Requirements for Teacher Certification." *Clearing House* 36: 158-60; November 1961.

BARRETT, JOHN A. *The Development of Certification Practices in New Mexico and the Study of the Relationship Between These Practices and Improved Educational Levels in the State.* Doctor's thesis. Albuquerque: University of New Mexico, 1961. 307 pp. Abstract: *Dissertation Abstracts* 22: 3050-51; No. 9, 1962.

BARTELS, MARTIN H. "The Index of Teacher Demand for 1958." *Journal of Teacher Education* 11: 15-19; March 1960.

BUSH, ROBERT N. "Responsibility for Professional Performance." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 271-76.

COREY, ARTHUR F. "The Role of the Professions in Licensure." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 136-41.

DAVIES, DON. "Professional Standards in Teaching: Moving from Ideas to Action." Abstract of address before the 1962 National Commission on Teacher Education and Professional Standards Regional Conferences. *Journal of Teacher Education* 13: 191-202; June 1962.

FOLSOM, MARION B. "Teacher Salaries and Shortages." *Phi Delta Kappan* 39: 341; April 1958.

FOX, RAYMOND B. "Factors Influencing the Career Choice of Prospective Teachers." *Journal of Teacher Education* 12: 427-32; December 1961.

GREENE, JAY E. "Qualifying Examinations plus Institutional Recommendation for Teacher Certification." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 248-52.

HARRIS, WILBUR S. *Factors Influencing the Withdrawal of Teachers from the Ohio Public School System.* Doctor's thesis. Bloomington: Indiana University, 1957. 322 pp. Abstract: *Dissertation Abstracts* 19: 274-75; No. 2, 1958.

HILLS, JOHN R., and DOWNS, HARRY S. "College Students' Attitudes Toward Teaching as a Profession." *Journal of Teacher Education* 13: 396-401; December 1962.

Hodenfield, G. K., and Stinnett, T. M. *The Education of Teachers: Conflict and Consensus.* Englewood Cliffs, N.J.: Prentice-Hall, 1961. 177 pp.

Lambert, Sam M. "Angry Young Men in Teaching." *NEA Journal* 52: 17-20; February 1963.

Lander, Richard. "What NCATE Means to the Teacher." *NEA Journal* 52: 30-31; February 1963.

Latimer, John F. "New Directions in Teacher Education."*Journal of Teacher Education* 14: 108-12; March 1963.

Lieberman, Myron. "Considerations Favoring National Certification of Teachers." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 200-10.

Lindsey, Margaret, editor. *New Horizons for the Teaching Profession.* Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. 243 pp.

McKean, Roland N., and Kershaw, Joseph A. "A Solution for Teacher Shortages." *Saturday Review* 45: 62-64, 75-76; November 17, 1962.

Maul, Ray C. "A Changing Trend in Teacher Supply." *NEA Journal* 49: 37; October 1960. (a)

Maul, Ray C. "Staff—Supply and Demand." *Encyclopedia of Educational Research.* Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1378-82. (b)

Maul, Ray C. "The Teacher Shortage Persists." *American School Board Journal* 140: 17-19, 47; May 1960. (c)

Maul, Ray C. "Can the Teacher Shortage Be Solved?" *American School Board Journal* 142: 17-19; May 1961.

National Education Association, Research Division. "The Teacher Shortage Analyzed." *NEA Research Bulletin* 38: 68-74; October 1960. (a)

National Education Association, Research Division. *Teacher Supply and Demand in Public Schools, 1960.* Research Report 1960-R7. Washington, D.C.: the Association, April 1960. 48 pp. (b)

National Education Association, Research Division. "Ten Years of Population Growth." *NEA Research Bulletin* 39: 104-106; December 1961.

National Education Association, Research Division. *Estimates of School Statistics, 1962-63.* Research Report 1962-R13. Washington, D.C.: the Association, December 1962. 32 pp. (a)

National Education Association, Research Division. "Teacher Supply and Demand in Public Schools." *NEA Research Bulletin* 40: 93-95; October 1962. (b)

National Education Association, Research Division. *Teacher Supply and Demand in Public Schools, 1962.* Research Report 1962-R8. Washington, D.C.: the Association, April 1962. 55 pp. (c)

National Education Association, Research Division. "Interesting Facts and Figures on American Education." *NEA Research Bulletin* 41: 3-9; February 1963. (a)

National Education Association, Research Division. *Rankings of the States, 1963.* Research Report 1963-R1. Washington, D.C.: the Association, January 1963. 68 pp. (b)

Osborn, Wayland W., and Maul, Ray C. "New Emphasis Needed in Teacher Recruitment." *Midland Schools* 75: 20-21; March 1961.

Phi Delta Kappan. "Keeping Abreast in Education." *Phi Delta Kappan* 44: 353-56; April 1963.

Richards, Rachel. "Prospective Students' Attitudes Toward Teaching." *Journal of Teacher Education* 11: 375-80; September 1960.

Seltzer, Richard W. "Why Don't More Students Prepare for Elementary School Teaching?" *Journal of Teacher Education* 13: 422-27; December 1962.

Shriver, Sargent. "When Peace Corps Teachers Return." *NEA Journal* 52: 13-14, 75; March 1963.

Slaughter, Eugene E. "The Use of Examinations for State Certification of Teachers." *The Education of Teachers: Certification.* Official Report of the San Diego Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. pp. 240-47.

Stiles, Lindley J. "New Roads to Certification." *NEA Journal* 48: 27; April 1959.

Stone, James C. *California's Commitment to Public Education.* New York: Thomas Y. Crowell Co., 1961. 193 pp.

TINK, ALBERT K. *Factors Related to Students Choosing or Not Choosing Teaching as a Vocation.* Doctor's thesis. Evanston, Ill.: Northwestern University, 1960. 250 pp. Abstract: *Dissertation Abstracts* 21: 1132-33; No. 5, 1960.

TUTTLE, DONALD R. "Editorial Comments: The Built-In Imbalance in American Education." *Journal of Teacher Education* 14: 3-4; March 1963.

TYLER, RALPH W., and MILLER, RICHARD I. "Social Forces and Trends." *NEA Journal* 51: 26-28; September 1962.

VAN PELT, RICHARD. *A Determination of the Practices and Procedures To Be Used in the Recruitment of Young People for Teacher Education Programs.* Doctor's thesis. Lincoln: University of Nebraska Teachers College, 1958. 105 pp. Abstract: *Dissertation Abstracts* 18: 1734; No. 5, 1958.

WALKER, RUSSELL L. *Factors Within the School Systems of Oklahoma Which Cause Teachers To Leave the Profession.* Doctor's thesis. Fayetteville: University of Arkansas, 1958. 204 pp. Abstract: *Dissertation Abstracts* 18: 1694; No. 5, 1958.

WOELLNER, ELIZABETH H., and WOOD, M. AURILLA. *Requirements for Certification, 1962-63.* Twenty-seventh edition. Chicago: University of Chicago Press, 1962. 143 pp.

YEAGER, HERMAN V. "Profession, Anyone?" *Journal of Teacher Education* 11: 460-63; December 1960.

CHAPTER II

Recruitment and Selection of Prospective Elementary and Secondary School Teachers

GLENN W. DURFLINGER

One of the best statements of the point of view in the teaching profession regarding the recruitment, selection, and guidance of prospective new members is Chapter VII of *New Horizons for the Teaching Profession,* a report of a project sponsored by the National Commission on Teacher Education and Professional Standards (Lindsey, 1961). Titled "Identification, Selective Admission, and Retention in Teacher Education," the chapter reflects the thesis that it is the responsibility of the teaching profession to determine what is required in training a successful teacher in terms of basic student characteristics before admission to the program, in recruiting and selecting candidates on the basis of this knowledge, and in exerting reasonable efforts to retain successful teachers within the active professional ranks. This point of view sets the stage for the research reviewed in this chapter of the Review.

Recruitment

The most important problem pertaining to recruitment is obtaining enough qualified teachers for the ever-increasing multitude of children in the nation's schools. School populations are expected to increase by approximately 1 million each year for at least a decade. This increase requires an annual addition of nearly 50,000 teachers. Since approximately 150,000 teachers retire from or leave the profession each year, a constant annual need for 200,000 additional teachers exists. Furthermore, there are 90,000-100,000 teachers with substandard teaching credentials—teachers who need to acquire additional training as soon as possible. Research on recruitment problems is desperately needed.

One of the most comprehensive surveys of recruitment was made by Page (1959), who reported on the major state and national organizations for teacher recruitment. This study involved a state-by-state analysis of the efforts expended by the National Congress of Parents and Teachers, the Student National Education Association, the Future Teachers of America (FTA), and other facilities to encourage young people to prepare for and to enter the teaching profession. Financial aid programs available to students were also listed. Such information as was revealed in this survey should be in the hands of high school and college counselors. In view of the urgent need in the nation's schools for reduction in the number of overcrowded classrooms, of restricted programs, and of limited curriculums, a survey such as that by Page should be conducted regularly to reveal the extent of progress in recruitment efforts.

Factors in Decision Making

A number of studies sought answers to questions tangentially related to recruitment. In a questionnaire study, Fox (1961) sought the factors underlying the decision to become a teacher. The following characteristics were found to be significantly related to the decision: (a) desire to work with children or adolescents, (b) liking for a particular subject, (c) opportunity to leave the profession and to return to it later, (d) increasing salary trend, (e) scores on inventories of vocational interest, (f) possibility of using teaching as a steppingstone to another career, and (g) membership in an FTA group. Factors b, d, e, and f were associated more strongly with the decisions of the men than with those of the women; and factors a, c, and g were more highly related to the decisions of the women than to those of the men.

Tink (1960) studied the factors related to the choice of teaching as a vocation. The responses from 914 high school students, 389 college students, and 132 teachers indicated that four influential factors in choosing teaching were (a) teaching is interesting and satisfying; (b) it is challenging and imaginative; (c) it offers opportunities for working with a favorite academic subject; and (d) it affords an enjoyable working relationship with pupils. It also was noted that experiencing harmonious relationships with people was very important. Reasons for not choosing teaching were (a) having a different vocational interest, (b) perceiving existence of unfavorable teaching conditions, and (c) feeling personal inadequacies.

Age at Making Decision

Tink (1960) found that women decided earlier than men to enter teaching. This conclusion was also reached by Fox (1961) and by Mason (1961) in a study at the U.S. Office of Education. There was only a slight discrepancy among the studies concerning the age at which the decision to enter teaching had been made. Tink also reported that one-third of the 389 college students in his sample decided to be teachers before they entered the tenth grade. The U.S. Office of Education survey revealed that 21 percent of the women and 7 percent of the men decided on a teaching career before the eleventh grade (Mason, 1961). Fox (1961) reported that about 80 percent of the women in preparation programs for teaching as well as a high proportion of prospective elementary teachers of both sexes had decided favorably about teaching before high school graduation, whereas only about 30 percent of the men and 50 percent of the prospective candidates for secondary school teaching had done so before high school graduation.

In a questionnaire study at the Ohio State University, Riccio (1961) found somewhat the same results. Of the 488 freshmen entering the teacher preparation program, the average ages at which the male and female candidates had made their decision were, respectively, 17.05 years and 15.08 years.

Two studies revealed that individuals sampled perceived that teaching offers the opportunity to leave the profession and to return to it later. Fox (1961) found that factor to be related statistically to decision making. In the U.S. Office of Education study by Mason (1961), more than four out of five of the women in their first year of teaching said that they would want to return from homemaking to teaching at some future time. The opportunity in teaching to take time off for homemaking and for children is potentially important in recruitment.

Murphy (1962) studied graduates of the University of Georgia to learn which programs most often led toward teaching. Students with a bachelor of science in education and those with majors in business education, elementary education, and physical education more often went into teaching than those with majors in agricultural education, English, or social sciences. Men were more likely than women, and younger students were more likely than older students, to change their minds about entering teaching. He also noted that the minimum entrance requirements for the University of Georgia served to select students with sufficiently high academic ability to teach successfully. The secondary school background and the freshman placement examination were not definitive factors relative to who would enter teaching.

Dutton and Keislar (1961) determined that high school students did not possess a significant amount of correct information concerning the teaching profession. They also found that boys thought teaching would be less remunerative than did girls. Students from an upper socioeconomic group believed that their parents would be unhappy about their choice of teaching as a career, although students from the middle and lower socioeconomic groups felt that their parents would be pleased. The authors concluded that attitudes toward teaching were formed long before high school graduation.

In their investigation of the attitudes of college students toward teaching, Hills and Downs (1962) concluded that men were opposed to entering the teaching profession because of limited income and restricted opportunity for advancement—a conclusion essentially in agreement with the one reached in the study by Dutton and Keislar (1961).

Selection

In the area of selection of students to pursue a teacher preparation program, research efforts either must rely upon previous investigations in the measurement and prediction of teacher effectiveness or must make some unique contribution to knowledge in this area.

In the follow-up of Stout's (1957a, b) study of selection procedures, Edson and Davies (1960) attempted to obtain information concerning both nature and influences of changes taking place in selection procedures between 1953 and 1958 in Minnesota colleges and universities. They found that there was an increase in the use of records and measures of academic achievement and in employment of indexes of competence in English usage as bases for selection. There was also an increase in the use of faculty committees in the selection of students. Colleges with relatively low or moderate selectivity standards were sending more teachers into the schools than were the colleges with higher standards.

Rutherford (1961) conducted a questionnaire study to determine current practices in selective admission and in selective retention in certain teachers colleges. Altenhein (1962) described the guidance and selection program at Hunter College. Although there was little systematic inquiry into the characteristics of the total program in teacher training, considerable research was undertaken concerning elements or details of student characteristics in the selecting process.

One such study at the University of Southern California by Michael and others (1961) indicated some of the types of instruments employed in the selection process as well as the results of their use. They found that success in a teacher preparation program could be predicted to the degree that about 20-25 percent of the variance in the criterion variable was explained. The battery of tests consisted of the social studies, literature, science, fine arts, and mathematics subtests of the *Cooperative General Culture Test;* an English fundamentals test developed at the University; and the reading vocabulary, reading comprehension, arithmetic reasoning, and arithmetic fundamentals parts of the *California Achievement Tests.* When the results of a scale called Rating Sheet for Directed Teaching, the grade point averages in methods courses, the ages of students, and their classifications in college were added to the other measures cited, the percentage of predicted variance rose to above 40.

Criteria of Selection

In a questionnaire study of 48 universities and 132 colleges concerning the type and place in the students' program of the use of selective devices, Magee (1961) found that most institutions selected students (a) at the beginning of the upper division and (b) at entrance into student teaching.

Ninety-two percent of the institutions imposed some form of official hurdle at the upper division level, and 55.6 percent required formal admittance to upper division professional courses. A grade of C was generally adequate for admission to upper division or to professional courses, and only 3.3 percent reported a grade point standard of 2.5 or greater. Only 35 institutions demanded a grade point average higher than 2.0. Moreover, these requirements, which were applied primarily to secondary teaching candidates, were imposed variously upon (a) total work, (b) professional courses, and (c) major courses.

Restricting his research to 97 elementary school student teachers, Darrow (1962) found a multiple R of .33 between the criterion of a teacher rating scale and the combination of the *Minnesota Teacher Attitude Inventory* (*MTAI*), *Iowa Tests of Educational Development*, *American Council on Education Psychological Examination for College Freshmen*, total college point hour ratio, and high school grade average. He recommended continued use of the first three of these measures in the selection of prospective teacher candidates. Although the validity coefficient of the composite of predictors was not high enough to convince one that the instruments employed were the most effective ones, the relatively low R might actually be the result of the use of a teacher rating scale of low reliability and of questionable validity. Rabinowitz and Mitzel (1961) contended that, if the combined predictors employed resulted in a multiple R no higher than .40 with a reliable measure of teaching success, one was screening out too many potentially successful teachers and leaving in too many unfit ones.

Another study with somewhat negative results was that of Robinson (1962). On the basis of the criterion of going into teaching, he found that the predictive validity of such instruments as the *MTAI*, the *Guilford-Zimmerman Temperament Survey*, and the *Nelson-Denny Reading Test* was so low that he recommended against their use in accepting or rejecting students for the teacher training program. Fielder (1960) obtained similar results when he employed this same criterion, i.e., entry or nonentry into classroom teaching. He also employed the criterion of continuing or not continuing in the teacher preparation program.

Academic Competence

Several studies were concerned with the academic quality of prospective teachers found in colleges. Mitzel and Dubnick (1961) summarized several types of evidence which they declared to be inconclusive. Using the *Ohio State University Psychological Test*, *Iowa High School Content Examination*, and a locally constructed mathematics examination, Snider and Long (1961) compared majors in business education, engineering, pharmacy, and education. Prospective teachers did as well as the business education majors and the engineering majors, but did not reach the level of the pharmacy students in these tests. The authors suggested that, since measures of academic aptitude were available, they would be employed as the chief or only screening device unless more valid ones should be developed which would measure other characteristics. Cummings (1960) reported that students in the College of Education at the University of Alabama did better on tests of intelligence than did nonvocationally oriented students. However, the female sex far outnumbered the male sex in numbers in this college and in other vocationally oriented programs. Moreover, the women did somewhat better than the men did on the tests of intelligence at this university.

On the other hand, Vertein (1961) studied a group of students from Wisconsin State College and Institute of Technology at Platteville who were preparing to be teachers and found them to have an average academic aptitude below that of the students not planning to teach. Both groups were below the University of Wisconsin academic average. However, most of them came from a restricted geographic locale, namely, southwestern and rural Wisconsin.

There is a considerable list of concomitant sociological factors of current professional interest and significance to the topic of the academic or intellectual aptitude of teachers in training. An important study by Statler (1959) included 257 students at the Ohio State University who were beginning their professional course sequence. Nearly 60 percent of the fathers were employed in business, in production, or in agriculture. Half of their families had annual incomes below $7,500—a figure which was 2.7 below the national average. Fifty-six percent were graduated from a high school of 150 students or fewer. Eighty-five percent of the men and 55 percent of the women transferred from another major into a teaching major. Constituting 38 percent of the total sample, men preferred to teach in science or mathematics; women preferred language or social studies.

In short, there was inconclusive and conflicting evidence concerning whether the students who were preparing to be teachers or who were hopeful of becoming teachers were of lower academic aptitude than were either the students of universities in general or the students preparing for other fields of life work. Moreover, the data were confused by other correlated factors, such as socioeconomic background of candidates and future economic prospects after entering the profession.

Making use of several college groups including those preparing to teach, Van Wagenen (1961) attempted to determine the relationship between academic achievement and personality factors revealed by the *Edwards Personal Preference Schedule* (*EPPS*), the *Guilford-Zimmerman Temperament Survey* (*GZTS*), the *Strong Vocational Interest Blank*, a scale concerning rigidity of attitudes regarding personal habits, a questionnaire about personal data and background, an anxiety scale, and an inventory of beliefs. For the most part, his data indicated that none of the scores of these instruments was consistently related to grade point average, regardless of the college group considered. Those planning to teach tended to rank higher in ascendance, sociability, and friendliness in the *GZTS* than did students in other groups.

Career Commitment and Selection

The individual preparing for a career in teaching spends considerable money and time, and the community invests heavily in education. As suggested previously, one of the important arguments to be employed in recruitment is the possibility of a teacher's leaving a position in the schools for an indefinite period of time with the opportunity of returning at a later date. This significant recruitment inducement is possibly a factor related to mortality in the teaching profession which may be responsible for some of the teacher shortage. Presumably it would be impossible or at least inadvisable to require adequately prepared teachers with credentials to remain in the profession. However, if selective techniques could be developed that would pick out the one teacher who would remain on the job for 40 years, the time spent in training and supervision as well as untold amounts of money could be saved.

There are few studies in the recruitment and selection of teachers which are concerned with turnover and termination—either permanent or temporary. An investigation by Murphy (1962) has already been mentioned under Recruitment. Another study, by Rabinowitz and Crawford (1960), dealt with the relationship between predictor variables obtained during student teaching and subsequent persistence in the profession. They found that 40 percent of the original group of 1,144 graduates responding to the questionnaire had taught continuously for five years. Six percent had not taught. High school teachers did have a much better record of persistence than did elementary school teachers. More than three-fourths of the men graduates were teaching five years later, but less than one-half of the women were doing so. The older graduates were much more likely to persist than were the younger ones. Those who did indicate satisfaction with their student teaching experience were more persistent than those who did not. A general summary would suggest that, in terms of the data found for the predictor variables studied, male students and women of 30 or more years of age tended to make a career of teaching.

Personality Characteristics

The *Minnesota Multiphasic Personality Inventory* (*MMPI*) is occasionally used to identify those personality traits of teachers which may be of some assistance in selection. Flanagan (1961) found that teachers considered to be superior did rank higher on the *MMPI* scales of hypochondriasis, depression, and the masculine side of the masculinity-femininity scale than did teachers not considered superior. Fielstra (1963) reported that the personal qualities of forcefulness, emotional balance, and wholesome philosophy of life were the only ones rated in student teaching which were discriminating between excellent and good first-year secondary school teachers.

Since the publication of the *MTAI*, there has been a considerable volume of research with this instrument upon attitudes of teachers and teacher education students. The *MTAI* measures only a limited number of teacher attitudes. There is a need for new instruments to furnish information about other attitudes.

Employing the *MTAI* as well as two scales concerning educational values, MacDonald and Doll (1961) obtained evidence that older, part-time students did possess more mature attitudes about teaching than did younger, full-time students. Darrow (1962) found the *MTAI* to be sufficiently related to the standing on a teacher rating scale that he recommended the use of the *MTAI* in selection of students in teacher education.

Vertein (1961) used the *MTAI* to compare one sample of students planning to teach with a second sample which he called an "academic population" and found that the former group excelled the academic group by a significant amount. Scott and Brinkley (1960) studied the improvement made in the *MTAI* scores by student candidates during a semester of practice teaching. Those whose initial scores were inferior to those of the supervising teachers improved significantly, whereas those whose initial scores excelled those of the supervising teacher failed to make statistically significant improvement. They concluded that these results indicate a lack of either predictive or concurrent validity in the *MTAI*.

Popham and Trimble (1960) also checked upon the validity of the *MTAI* in their research to ascertain whether it could differentiate between groups of teachers rated successful or unsuccessful by administrative superiors. They stated that, since the inventory did differentiate between these groups at the .01 level, it should have considerable validity in selecting teacher education candidates.

Sheldon, Coale, and Copple (1959) assessed the degree of difference in certain personality traits of students aspiring to be teachers who ranked high or low in warmth or friendliness on prediction scales. They found that those who were high in these characteristics as compared with those who were low not only were significantly higher in intelligence and lower in authoritarianism but also expressed a significantly higher need for affiliation and a lower need for succorance. The *EPPS* and the *Thematic Apperception Test* (*TAT*) showed conflicting results in the needs of "warm" teachers for dominance and aggression as well as in their being characterized by these two traits. These conflicting findings suggested a lack of construct validity in the traits measured by either one, or by both, of the instruments. Conceivably, the traits may be parts of a larger personality configuration that is in need of further analysis.

Lang (1960) conducted a study which was more directly concerned with motives of students in deciding upon teaching at the elementary or secondary school level than were many other related investigations. He concluded that nurturance or "mothering" is significant in the choice of elementary level teaching and that achievement is relatively less important. For students choosing to teach in high school, the order of importance of these motives is reversed. Employing the *EPPS*, Southworth (1962) made a breakdown of the entire range of elementary teacher preparation into upper elementary and lower elementary levels and then studied the characteristic needs of teachers preparing for each level. He found that candidates preparing for the early or lower elementary level teaching were characterized by a greater need for abasement, affiliation, succorance, and nurturance; whereas the students preparing for upper elementary teaching revealed the need for achievement, aggression, and exhibition.

Garrison and Scott (1961) studied the personal needs of students who were preparing to teach at one or more levels within the span from kindergarten through high school. The students were classified into five teaching areas: (a) lower elementary, (b) upper elementary, (c) general secondary, (d) nongeneral secondary, and (e) special education. Those students planning to teach at the general and nongeneral secondary level were further divided according to the subject areas in which they were preparing to instruct. In this aspect of the research design, it was assumed that there may be differences in the personal needs measured according to the prospective teaching level and the subject to be taught. Although the analytic material was detailed and long, the following findings were suggestive: (a) The general secondary women did exhibit a significantly greater need for achievement than did women in either the elementary or the nongeneral secondary group. (b) The prospective teacher of lower elementary grades did manifest to a significantly greater degree the need for nurturance than did the upper elementary, general secondary, nongeneral secondary, and the prospective teachers of social studies, language arts, and business education.

In still another study involving use of the *EPPS*, Saltz (1960) investigated the difference between the stated needs of teachers and the stereotypes as indicated by nonteachers who expressed opinions of teachers' manifest needs.

From the results of the last five studies, all of which used the *EPPS*, it might be said in summary that the manifest needs of prospective teachers differed according to their proposed teaching level and to some extent according to the expected teaching area in high school. Elementary teachers in general did exhibit a significantly greater need for nurturance, succorance, affiliation, change, and abasement than did high school teachers. High school teachers in turn did show a significantly greater need for achievement, aggression, exhibition, and order than did elementary school teachers.

In the Saltz (1960) study, another valuable contribution was the hypothesis that the nonteacher's opinion of the teacher or of teachers in general—the stereotype—may tend to mould the asserted needs of student-teaching candidates toward this stereotype of the teacher. Taking a contrary point of view, Cook, Linden, and McKay (1961) suggested that students planning to become teachers possess within their personalities certain basic need orientations which are satisfied best through participation in teaching and in activities related to education. Although further research upon this issue is called for, the point made by Saltz is highly pertinent. If the stereotype of the teacher in the minds of potential teachers appears to be more or less repulsive, then recruitment of students for the program in teacher training becomes more difficult.

These few findings from the studies reviewed are sufficient to show the detail in which the results may be described. Although teachers in general may be characterized by certain descriptive terms, within the profession there appear to exist significant differences in traits which the teachers of each sex possess relative to both subject matter and grade level taught.

In their second study involving a sample of students who chose teaching as a career, Garrison and Scott (1962) considered the relationship between (a) such background factors as age, marital status, class standing in college, and years one planned to engage in teaching and (b) specific psychological needs as measured by the *EPPS*. Of the 15 psychological needs, 14 varied according to one or more background characteristics. Of the 12 characteristics, only 2 were unrelated to any of the 15 psychological needs.

Factor Studies

From the results of a number of factor-analytic studies, a partial basis is furnished for the identification of the characteristics which may be employed to select candidates for programs in teacher education. In one such study, Cook, Linden, and McKay (1961) analyzed the *EPPS* and the *GZTS* and identified the following six factors: (a) docility, (b) dependency, (c) authoritarianism, (d) compulsive conformity, (e) introversion-extroversion, and (f) avoidance. Using the technique of having faculty members in a department of education check a list of statements descriptive of a good teacher, Myers (1960) isolated four factors through a statistical analysis: (a) general acceptability, (b) readiness to learn, (c) social sensitivity, and (d) impulse control.

Mitchell and Pierce-Jones (1960) carried out a factor analysis of scores on scales of the *California Psychological Inventory* (*CPI*) that had been administered to 258 university students. The investigation listed four major factors provisionally identified as (a) adjustment by social conformity, (b) social poise, (c) superego strength, and (d) capacity for independent thought and action. The results of this study pointed to a possible rearrangement of some of the 18 subtests of the *CPI*. The authors suggested that the 18 *CPI* scales did not represent 18 independent personality dimensions.

Perry (1962) carried out a factor analysis of the intercorrelations of 43 predictor variables and of the ratings of four judges of the teaching performance of secondary school student teachers. The factors were identified as (a) intellectual ability, (b) social adjustment, (c) professional enthusiasm, and (d) speaking ability. There were few conclusions, if any, which could be derived from this group of factor studies. The populations employed in the testing were dissimilar, and the instruments also differed.

A word of caution should be given to anyone attempting to do research with the factor-analysis research tool—especially with personality scales. *Before commencing a factor analysis project, one should review the conditions that influence the outcomes of factor analysis.* Only one source of error or misinterpretation will be mentioned here. Spurious correlations arise with the use of scores that are experimentally dependent, e.g., in personality tests in which the scores are derived from overlapping keys.

A study by Jackson and Moscovici (1963) concerning the personality characteristics of students identified prior to their assuming the role of the teacher was noteworthy because of the measuring scales employed. They used three paper-and-pencil tests. The subjects (a) drew a picture of a teacher, (b) responded to a word association test in which the words might be related to teaching or to the school, and (c) participated in a word completion test similarly related to teaching and the school. Results of this research indicated that trainees at the beginning of their preparation had begun to identify with the teacher role. They also revealed a type of conflict or uneasiness with the authority role that they would soon assume as well as the desire to maintain a pleasant, protective, nonauthoritarian school atmosphere. Minimizing the significance of the conflict aspects of the study and maximizing the novelty of the approach employed in the research endeavor, one can hope for more studies of this type.

Highly recommended is one general reference by Getzels and Jackson (1963) which is a nearly complete bibliography of structure concerned with the personality and characteristics of the teacher. The student of recruitment and of selection in programs of teacher training will find the reading of this chapter especially rewarding.

Conclusion

Although noteworthy strides have been made in the improvement of research designs and in the introduction of extensive statistical analyses of data related to problems of recruitment and selection, one cannot help but be impressed by the lack of replication of investigations. Since no two researchers seem ever to use the same test batteries or to make use of populations that could be judged comparable, it is difficult to ascertain whether any genuine direction in the findings may be detected. It would be helpful if large-scale cooperative studies could be undertaken in which attempts are made to develop a systematic body of knowledge.

Bibliography

ALTENHEIN, MARGARETE R. "Guidance and Selection in a Teacher Education Program." *Journal of Educational Sociology* 36: 70-78; October 1962.

COOK, DESMOND L.; LINDEN, JAMES D.; and MCKAY, HARRISON E. "A Factor Analysis of Teacher Trainee Responses to Selected Personality Inventories." *Educational and Psychological Measurement* 21: 865-72; Winter 1961.

CUMMINGS, JAMES F. *A Study of Certain Characteristics of the Student Body in the College of Education at the University of Alabama, 1955-59.* Doctor's thesis. Tuscaloosa: University of Alabama, 1960. 192 pp. Abstract: *Dissertation Abstracts* 21: 1864; No. 7, 1961.

DARROW, HARRIETT D. "The Relationships of Certain Factors to Performance of Elementary Student Teachers with Contrasting Success Records in Student Teaching." *Teachers College Journal* 33: 95-98; January 1962.

DUTTON, WILBUR H., and KEISLAR, EVAN R. "Attitudes Toward Teaching." *Journal of Teacher Education* 12: 165-71; June 1961.

EDSON, WILLIAM H., and DAVIES, DON. "Selectivity in Teacher Education." *Journal of Teacher Education* 11: 327-34; September 1960.

FIELDER, WILLIAM R. *A Study of Certain Antecedent Factors Associated with Perseverance in Training and Entry into Classroom Teaching.* Doctor's thesis. Stanford, Calif.: Stanford University, 1960. 144 pp. Abstract: *Dissertation Abstracts* 21: 97; No. 1, 1960.

FIELSTRA, CLARENCE. "Discriminative and Predictive Values of Ratings Given to UCLA Student Teachers on the Secondary School Level." *California Journal of Educational Research* 14: 11-18; January 1963.

FLANAGAN, CARROLL E. "A Study of the Relationship of Scores on the MMPI to Success in Teaching as Indicated by Supervisory Ratings." *Journal of Experimental Education* 29: 329-54; June 1961.

FOX, RAYMOND B. "Factors Influencing the Career Choice of Prospective Teachers." *Journal of Teacher Education* 12: 427-32; December 1961.

GARRISON, KARL C., and SCOTT, MARY H. "A Comparison of the Personal Needs of College Students Preparing To Teach in Different Teaching Areas." *Educational and Psychological Measurement* 21: 955-64; Winter 1961.

GARRISON, KARL C., and SCOTT, MARY H. "The Relationship of Selected Personal Characteristics to the Needs of College Students Preparing To Teach." *Educational and Psychological Measurement* 22: 753-58; Winter 1962.

GETZELS, J. W., and JACKSON, P. W. "The Teacher's Personality and Characteristics." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 11, pp. 506-82.

HILLS, JOHN R., and DOWNS, HARRY S. "College Students' Attitudes Toward Teaching as a Profession." *Journal of Teacher Education* 13: 396-401; December 1962.

JACKSON, PHILIP W., and MOSCOVICI, FELA. "The Teacher-To-Be: A Study of Embryonic Identification with a Professional Role." *School Review* 71: 41-65; Spring 1963.

LANG, GERHARD. "Motives in Selecting Elementary and Secondary School Teaching." *Journal of Experimental Education* 29: 101-104; September 1960.

LINDSEY, MARGARET, editor. *New Horizons for the Teaching Profession.* Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. 243 pp.

MACDONALD, JAMES B., and DOLL, RONALD C. "Who Is Ready for Teacher Education?" *Journal of Educational Sociology* 35: 123-27; November 1961.

MAGEE, ROBERT M. "Admission-Retention in Teacher Education." *Journal of Teacher Education* 12: 81-85; March 1961.

MASON, WARD S. *The Beginning Teacher: Status and Career Orientations.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 664. Washington, D.C.: Government Printing Office, 1961. 196 pp.

MICHAEL, WILLIAM B., and OTHERS. "The Prediction of Success in Selected Courses in a Teacher Training Program from Scores in Achievement Tests and from Ratings on a Scale of Directed Teaching Performance." *Educational and Psychological Measurement* 21: 995-99; Winter 1961.

MITCHELL, JAMES V., JR., and PIERCE-JONES, JOHN. "A Factor Analysis of Gough's California Psychological Inventory." *Journal of Consulting Psychology* 24: 453-56; October 1960.

MITZEL, HAROLD E., and DUBNICK, LESTER. "The Relative Scholastic Ability of Prospective Teachers." *Journal of Teacher Education* 12: 73-80; March 1961.

MURPHY, MARY O. "A Study of Graduates of the University of Georgia Who Are Certified To Teach with Respect to Entrance into the Teaching Profession." *Journal of Experimental Education* 31: 55-64; September 1962.

MYERS, ROGER A. "Faculty Perceptions of Student Success in Teacher Education." *Journal of Educational Psychology* 51: 235-39; August 1960.

PAGE, M. JEROME. "Organization for Teacher Recruitment." *Journal of Teacher Education* 10: 134-56; June 1959.

PERRY, JAMES O. *A Study of a Selective Set of Criteria for Determining Success in Secondary Student Teaching at Texas Southern University.* Doctor's thesis. Austin: University of Texas, 1962. 243 pp. Abstract: *Dissertation Abstracts* 23: 1617; No. 5, 1962.

POPHAM, W. JAMES, and TRIMBLE, ROBERT R. "The Minnesota Teacher Attitude Inventory as an Index of General Teaching Competence." *Educational and Psychological Measurement* 20: 509-12; Autumn 1960.

RABINOWITZ, WILLIAM, and CRAWFORD, KAY E. "A Study of Teachers' Careers." *School Review* 68: 377-99; Winter 1960.

RABINOWITZ, WILLIAM, and MITZEL, HAROLD E. "Some Observations on the Selection of Students for Teacher Education Programs." *Journal of Teacher Education* 12: 157-64; June 1961.

RICCIO, ANTHONY C. "When Do They Decide To Teach?" *Peabody Journal of Education* 39: 94-95; September 1961.

ROBINSON, WILLIS. *A Validity Study of the Testing Program for the Selection of Students for Teacher Education.* Doctor's thesis. West Lafayette, Ind.: Purdue University, 1962. 75 pp. Abstract: *Dissertation Abstracts* 23: 2812; No. 8, 1963.

RUTHERFORD, ROBERT D. *Recommendations for Improved Selection and Retention Processes for Elementary Teacher Candidates in Selected State Teachers Colleges.* Doctor's thesis. Denver, Colo.: University of Denver, 1961. 229 pp. Abstract: *Dissertation Abstracts* 22: 3939-40; No. 11, 1962.

SALTZ, JOANNE W. "Teacher Stereotype—Liability in Recruiting?" *School Review* 68: 105-11; Spring 1960.

SCOTT, OWEN, and BRINKLEY, STERLING G. "Attitude Changes of Student Teachers and the Validity of the Minnesota Teacher Attitude Inventory." *Journal of Educational Psychology* 51: 76-81; April 1960.

SHELDON, M. STEPHEN; COALE, JACK M.; and COPPLE, ROCKNE. "Concurrent Validity of the 'Warm Teacher Scales.'" *Journal of Educational Psychology* 50: 37-40; February 1959.

SNIDER, GLENN R., and LONG, DELBERT. "Are Teacher Education Programs Attracting Academically Able Students?" *Journal of Teacher Education* 12: 407-11; December 1961.

SOUTHWORTH, HORTON C. *A Study of Certain Personality and Value Differences in Teacher Education Majors Preferring Early and Later Elementary Teaching Levels.* Doctor's thesis. East Lansing: Michigan State University, 1962. 114 pp. Abstract: *Dissertation Abstracts* 23: 1284; No. 4, 1962.

STATLER, ELLSWORTH S. "Characteristics of Students in a Teacher-Education Course." *Educational Research Bulletin* 38: 151-58, 168; September 1959.

STOUT, RUTH A. "Selective Admissions and Retention Practices in Teacher Education." *Journal of Teacher Education* 8: 299-317, September; 422-32, December 1957. (a)

STOUT, RUTH A. *A Study of Admissions and Retention Practices in College and University Programs of Teacher Education.* Doctor's thesis. Minneapolis: University of Minnesota, 1957. 348 pp. Abstract: *Dissertation Abstracts* 17: 1953-54; No. 9, 1957. (b)

TINK, ALBERT K. *Factors Related to Students Choosing or Not Choosing Teaching as a Vocation*. Doctor's thesis. Evanston, Ill.: Northwestern University, 1960. 250 pp. Abstract: *Dissertation Abstracts* 21: 1132-33; No. 5, 1960.

VAN WAGENEN, DONALD R. *The Relation of Selected Non-Intellectual Factors to Academic Achievement in Several College Groups*. Doctor's thesis. Syracuse, N.Y.: Syracuse University, 1961. 179 pp. Abstract: *Dissertation Abstracts* 23: 539; No. 2, 1962.

VERTEIN, LESTER D. "A Study of the Personal-Social and Intellectual Characteristics of a Group of State College Students Preparing To Teach." *Journal of Experimental Education* 30: 159-92; December 1961.

Additional References

DOANE, KENNETH R., and SCANLAN, WILLIAM J. "Saint Paul, Minnesota, Teacher Recruitment Project Accounts for the Third Year of College and School Experiences." *Bulletin of the National Association of Secondary-School Principals* 45: 119-45; January 1961.

EDSON, WILLIAM H. "Selecting Students for the College of Education at the University of Minnesota." *Journal of Teacher Education* 14: 51-56; March 1963.

RICHARDS, RACHEL. "Prospective Students' Attitudes Toward Teaching." *Journal of Teacher Education* 11: 375-80; September 1960.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE, OFFICE OF EDUCATION. *Annual Report, 1961*. Washington, D.C.: Government Printing Office, 1962. 418 pp.

CHAPTER III

Pre-Service and In-Service Education of Teachers

HAROLD E. REYNARD

PUBLICATIONS by both professional educators and laymen highlighted the current intense interest in the preparation of teachers. Unfortunately, indications of bold experimentation either in curriculum offerings or in methods used in teacher education could not be found. Evidence of research in teacher education, as reported in the literature between January 1958 and March 1963, is abundant in relation to field experience, measures of teachers' attitudes, and programs designed primarily for liberal arts college graduates. Little evidence of experimentation with the total program of teacher education was found for this period.

Recent General Developments

Increasing pressures to extend requirements and to improve the quality of all three aspects of teacher education—general education, field of specialization, and professional education—apparently have overtaxed the limits of four-year programs. Additional obstacles to program changes were increasing enrollments, intensive demands for more teachers by the public schools, the changing nature of many institutions from a single-purpose to a multipurpose curriculum, and lack of adequate resources and facilities.

Meyer (1960) studied the extent of conversion among single-purpose teachers colleges to multipurpose institutions. He found that in the conversion the institutions extended their programs both horizontally and vertically. Meyer suggested the possibility that many colleges that have expanded may be tending toward overspecialization and vocationally oriented programs. Criticism of teacher education also may have been a factor in the apparent piecemeal approach to program changes. In a survey of books and magazines for the period from January 1, 1948, to July 1, 1958, Popham and Greenberg (1958) found that such criticism increased sharply during 1957-58. Their data showed that college and university personnel, other than professional educators, provided the largest number of criticisms of teacher education programs.

It was clearly indicated in the writings of Andrews and Palmer (1958), Hodenfield and Stinnett (1961), Lindsey (1961), and Trump (1962) that the preparation necessary for teachers cannot be provided in a four-year period. A serious effort to develop an improved four-calendar-year program for preparing elementary teachers was made in the Yale-Fairfield Study, which was reported by Howell and others (1960). Longer than four academic school years, this program was an attempt to overcome two major problems in teacher education: too little professional laboratory experience and too little general education. This program, which utilized three summers and four school years, provided extensive field experience with little sacrifice to the general education program.

Pre-Service Programs

Research concerning the effectiveness of teacher education programs has been confined largely to questionnaire or opinionnaire surveys. Variables in such domains as personality, curriculum, and methods continue to defy satisfactory control and measurement. This lack of control and of precision in measurement may explain the reluctance for experimentation in teacher education. Clearly defined criterion measures of the competencies and qualities that are needed to be an effective teacher should hasten the development of the satisfactory planning of programs and their evaluation. Fishburn (1962) found significant differences between teachers' and administrators' concepts of the teacher's role. Follow-up studies of teachers, except for surveys of their opinions about the programs, were conspicuously absent from the literature of the past five years.

Of some relevance to the pre-service program has been the work of campus-wide committees concerned with teacher training programs. Goodlad and Boozer (1961) pointed out that the number of campus-wide committees representative of all faculty members who work with pre-service teachers has increased steadily since 1890. These writers posed several insightful questions concerning the need for and value of such committees.

Elementary School Teacher Education

To evaluate the program for the preparation of elementary teachers at the University of Oregon, Piper (1960) surveyed 155 graduates and 118 principals of the schools in which the graduates had taught. It was determined that 42 percent of the graduates preferred both the same amount of work in each course and the same sequence. Many of the respondents recommended an increase in the number of hours in science. The graduates of this group who did not teach expressed more dissatisfaction with their preparation than those who were teaching.

Through the study of college catalogues, Andrews and Palmer (1958) examined the programs for the preparation of elementary teachers of 22 public and private institutions of seven different types in 19 states. Curriculum requirements for degree programs were studied to determine the number of credit hours devoted to (a) general education, (b) additional subject matter, (c) professionalized subject matter, (d) professional courses, and (e) electives. Credit hours required for general education represented, on an average, 25 to 33 percent of the total hours, and professional education represented between 20 and almost 40 percent of the work counting toward the degree. The number of hours allowed for electives ranged from zero to 40 hours with slightly less than 33 percent of the institutions permitting nearly one-third of the total program to be used for electives. Although the general education provided by these institutions was broad, it not only lacked depth but also provided a very limited base for graduate work. As determined by the survey, the amount of work in foreign languages, literature, and mathematics was either omitted or slighted in most programs.

Wiebe (1960) undertook a survey of elementary teacher education in small church-related liberal arts colleges and universities in California. Information was obtained from college catalogues, from interviews with college staff members, and from a questionnaire sampling of those graduates for the 1954, 1955, and 1956 school years who had taught in the California public schools. For the 15 colleges involved, 325 graduates returned questionnaires. Wiebe reported considerable variation among the schools in the type and in the amount of field experience required. He also indicated that systematic follow-up programs generally were not provided in these institutions.

Secondary School Teacher Education

Increasing public school enrollments, which resulted in many new high schools being built to house students in grades 10, 11, and 12, have directed attention toward the in-between groups—the junior high school students. Concern for the preparation of junior high school teachers prompted Ackerman (1960) to analyze curriculum offerings that had been designed especially for the preparation of these teachers. Through examination of the catalogues of 246 National Council for Accreditation of Teacher Education (NCATE) institutions, he classified 46 as not having provisions for junior high school teachers, 164 as providing separate courses, and only 36 as offering special curriculums for junior high school teachers. He concluded that there are not many institutions that provide the same attention to the preparation of junior high school teachers as to the training of elementary and high school teachers.

Cogan (1958) examined catalogues of 114 public and private colleges and universities which prepare secondary school teachers. One purpose of the study was to determine the quantitative relationship between professional credit hour and the total credit hour requirements for a bachelor's degree. Five classifications for the preparing institution were used: (a) public teachers colleges, (b) public liberal arts colleges, (c) private liberal arts colleges, (d) public universities, and (e) private universities. Cogan's interpretation of this study was that the requirements in professional education were nearly the same for all five types of institutions. On the basis of his findings, he estimated that 16 to 20 percent of the total course work for a degree was devoted to professional education.

Many undergraduate programs for the preparation of teachers were described in part five of the Official Report of the Kansas Conference of the National Education Association, National Commission on Teacher Education and Professional Standards (1959).

Strong interest in the subject matter preparation of the secondary school teacher was evident in the literature. In order to determine the quantity and type of specialized study of English language other than composition required for high school English teachers, Slothower (1959) sent a questionnaire to 395 English department heads, of whom 360 responded. It was determined that 71 percent of the institutions required one or more courses in the special study of the English language for teachers majoring in English. Of the 256 requiring special study, 77 percent required only one course. Prescribed work in grammar predominated most programs that were intended for the preparation of English teachers.

In a short article, Mayor (1962) introduced a group of nine reports concerning special projects that had been designed to improve the science and mathematics preparation of teachers. These projects were part of the Science Teaching Improvement Program of the American Association for the Advancement of Science supported by the Carnegie Corporation. In a historically oriented study, Schumaker (1959) determined that only slight changes occurred between 1920 and 1958 in the requirements for teachers who majored in teaching mathematics at the secondary level. Questionnaires were sent to 314 mathematics department heads, and course requirements for 140 institutional programs were examined. The data revealed that elementary mathematics through calculus, including solid geometry, differential equations, and theory of equations, represented the preparation required in the academic year of 1920-21. At that time, a minimum of 24 semester hours was the median requirement for major preparation. During the 1957-58 school year, the median number of semester hours required was 27. The mathematics educators' recommendations for teacher education courses were consistent from 1920 to 1958.

In a study of human relations in the classroom, the North Central Association of Colleges and Secondary Schools, Subcommittee on Human Relations in the Classroom (1962) reported little concern by college and university staff members and no organized effort for preparation of teachers for successful human relations with students, parents, and the community. Questionnaire returns from 1,108 professors from several disciplines representing 336 North Central institutions were examined to determine concern for and pre-service practices for preparing teachers to deal with children as persons, to work with parents of all kinds, and to understand community conditions that affect children's development. Although this survey showed that large numbers of professors valued this aspect of teacher preparation highly, they placed little emphasis on it in their courses. The North Central Association of Colleges and Secondary Schools, Subcommittee on Human Relations in the Classroom (1963) also surveyed 1,075 second-year secondary school teachers in order to determine not only the problems that teachers have in interpersonal and intergroup relations but also their perceptions of the curriculum experiences provided in their pre-service programs that were helpful in the understanding of human relations and in the development of appropriate behaviors. Responses to questionnaires showed that one-half of the teachers considered college courses of little aid in their preparation for dealing with human relations problems. However, those courses which they considered helpful in this area were psychology, student teaching, and the social sciences. For the teachers surveyed, the data revealed that in the North Central schools differences in social class were a source of more problems than were matters involving relations with minority groups. Information obtained from these studies should be useful for program planning at both pre-service and in-service education levels.

In-Service Programs

Cooperation between public schools and teacher education institutions for in-service education of teachers has been stimulated by the "add-on" fifth-year programs for liberal arts graduates. Tremendous impetus for in-service education has been supplied by institutes in subject matter fields. Support for both summer and year-round institutes in mathematics, science, languages, and guidance provided by the National Defense Education Act (NDEA), the National Science Foundation (NSF), and various foundations has provided greatly increased resources for in-service teachers. Many descriptions of institutions have appeared in the literature. However, little has been written concerning in-service programs of public schools or concerning cooperative efforts between teacher education institutions and public schools for the development of such programs.

McCreary (1960) reported that school superintendents and academic personnel in colleges attempted to determine the individual needs of elementary teachers at the time of graduation in order that an in-service individualized orientation program could be planned. A small sample of administrators and teacher candidates participated. An interview and a diagnostic check list, which was based on criteria furnished by the superintendents, were used to ascertain candidates' strengths and weaknesses. Although the results indicated that additional adaptations were necessary, the instrument and procedure were considered useful. This approach holds promise for planning in-service programs.

In a questionnaire study of 218 in-service teachers, Taylor (1961) studied their plans for additional formal education as well as their need for further preparation. Nearly 75 percent of this group intended to take additional college work. Slightly more than 55 percent had plans to complete a master's degree in the near future. Classroom control and student motivation were listed most frequently as the most important problems they faced.

In an earlier study of the extent and status of in-service education in 100 Indiana schools, Taylor (1959) found that wealth of a school district was a principal factor in the amount of in-service education provided by a school system. The larger, wealthier city school districts provided more in-service education than the smaller, poorer districts.

Designing workshops and institutes to serve best the continuing educational needs of teachers in the field confronts both public schools and teacher education institutions. Fowler (1960) described an NSF-supported summer institute in 1959 for the preparation of 45 elementary school science resource teachers. Each of the teachers had at least three years of classroom experience and was teaching during the year prior to the workshop. Increased achievement in general science as measured by the test selected and administered prior to and following the program was considered to be highly significant.

In a survey of 28 colleges and universities, Woronoff (1958) classified four patterns of teacher education programs for liberal arts graduates. In nine universities, courses were offered to permit candidates to meet only minimum state certification requirements. Eleven universities required a series of courses or a prescribed pattern to be completed before the institution recommended certification. Six institutions offered special seminars to provide supervised group experiences as well as supervised student teaching, with much of the work applicable to a master's degree. Only two institutions provided individualized programs leading both to a master's degree and to recommendation for certification.

New Media in Teacher Education

Opportunities exist for experimentation with two new media in teacher education: (a) programed instruction and (b) closed circuit television. Little was found in the literature since 1958 concerning the use of programed instruction in teacher education. Interest in the use of closed circuit television in pre-service programs has led to several experiments in its employment during the 1958-63 period. Rumford (1960) experimented with the use of closed circuit television in the teaching of elementary school methods. Two groups, one experimental group and one control group, of 15 pairs of students matched on several factors participated. Teacher-developed objective tests were given to measure differences in attainment. In achievement, the TV lectures were judged to be as successful as conventional instructional methods. Although it was found that television was as effective as direct observation for entire groups, it was less advantageous for high-ability students than for students of less ability, relative to measures obtained from the objective tests. Moreover, in terms of final grades and final examination scores received by students, television was considered to be as successful as other methods used. In relation to changes in attitude as measured by the *Minnesota Teacher Attitude Inventory* (*MTAI*) scores, television was as influential as conventional methods.

Thompson (1960) carried out an experiment with three methods of observation—direct classroom observation, observation by closed circuit television, and observation by films—to determine the relative values of each for changing selected student attitudes. He presented the following conclusions: (a) attitudes of the students toward children were not influenced to a significant degree by any of the three methods of observation; (b) direct observation was preferable to that based on television, and television observation was superior to film observation; and (c) differences in instructors appeared to influence student attitudes more than the method of observation. Other studies, which were listed in Thompson's bibliography, reported similar results concerning these three methods of teaching.

Professional Laboratory Experience

Professional laboratory experience seems to be the area least challenged in teacher education. Studies concerning the increased use of early direct experience with children prior to student teaching, as well as the intern experience that follows student teaching, appeared in the literature for the period of this REVIEW. Williams, Deever, and Flynn (1960) analyzed responses from 14 teacher education institutions in Oklahoma to determine the extent of and value placed on all types of professional laboratory experiences. Reports from these institutions clearly indicated that participation and observation were considered effective and essential for students prior to student teaching. Administrative arrangements for providing student teaching as well as other types of field experience in these institutions also were reported. In a related study, Levine (1961) indicated that student participants considered short-term experience during the freshman year, such as assisting in public school offices and in classrooms, to be worthwhile.

Investigators appeared to give more attention to student teaching than to other aspects of teacher education. Changes in attitudes and beliefs, the amount of time and supervision provided for student teaching, students' evaluations of their programs by questionnaires, and the selection and preparation of public school supervisors were the topics studied most frequently. From a critical analysis of selected research concerning student teaching, Young (1961) developed a set of guiding principles for planning and evaluating such programs. Inlow (1960) investigated the practices in 38 Midwestern teachers colleges and liberal arts colleges and universities in relation to the required and elective hours for student teaching. She reported that the mean number of hours for elementary student teaching and for secondary student teaching were, respectively, (a) 13.45 and 12.50 in the teachers colleges, and (b) 15.50 and 14.49 in the universities,

(c) 10.85 and 9.97 in the liberal arts colleges. Inlow also presented information concerning compensation to cooperating schools and teachers, selection and competencies of cooperating teachers, field experience provided prior to student teaching, and many other practices of the 38 institutions.

Through a study of 551 questionnaire responses, Jones (1960) found that full-time college supervisors were responsible for an average of 18 off-campus student teachers and that part-time supervisors were responsible for 15 off-campus student teachers. On the average, four student teachers were assigned each cooperating school. In addition, he found that college supervisors typically visited their student teachers three or four times during the teaching term.

Price (1961) studied the influence of public school supervising teachers on the attitudes and practices of student teachers. His findings indicated that student teachers adopted many of the practices of their public school supervisors during the student teaching period. He noted that students' attitudes changed in the direction of those held by their respective public school supervisors. McAulay (1960) observed six young teachers during their first year of teaching to determine which of the practices and materials they used were similar to those of their respective public school student teaching supervisors. During their student teaching, the six young trainees had been under the direction of one of three supervising teachers. The investigation concluded that the materials and techniques prepared and used during the period of student teaching were employed more extensively during the teacher's first year of teaching than were those materials and techniques developed in methods courses.

Many writers in teacher education stressed the need for programs aimed at the preparation of public school supervisors. Frantic searching for adequate numbers of schools to furnish facilities as well as for qualified supervisors to direct professional laboratory experience nearly has eliminated the opportunity either to select supervisory candidates or to provide them with special preparation and experience. Hayes (1960) surveyed 10 teacher education institutions to find what they were doing to prepare beginning supervising teachers. Subsequent to visiting the institutions, interviewing the faculty members, and studying the literature provided, he arrived at the following conclusions: (a) six institutions offered a special course in the pre-service program for the preparation of supervising teachers; (b) the most prevalent methods for orienting supervising teachers were by means of conferences and literature; and (c) there was no organized effort to have the supervising teachers become familiar with the teacher education program. With respect to provision for additional preparation of the experienced supervising teacher, Hayes presented the following information: (a) only two teacher education institutions had provided specific budgets or plans for financing in-service preparation programs; (b) group conferences and workshops commonly were used for improving supervision; (c) data sheets for student teachers were supplied to supervisors in all but one institution; and (d) only 5 of the 10 institutions attempted to bring supervisors and students together before the student began student teaching. He concluded that little had been done at the pre-service level to prepare supervisors and that lack of specifically assigned finances hindered the development of adequate in-service programs for experienced supervisors.

The Campus Laboratory School

Much study and thought have been devoted to the role of the campus laboratory school in teacher education. Inspection of the professional literature since 1958 indicated that use of laboratory schools for student teaching has been declining. However, in the opinion of several writers, growth of television in teacher education, increases in the amount and kinds of field experiences other than student teaching, and increased commitment to research have maintained the value of campus schools for both pre-service and in-service education of teachers. Despite these positive points, Ohles (1961) has described clearly the dilemma facing the laboratory school.

Hughes (1959) not only made a comprehensive study of laboratory schools but also examined their roles from a historical point of view. He visited 7 campus schools and received questionnaire returns from 29 others. Among the several findings determined from his visits to the campus schools, two were particularly relevant: (a) the role of campus schools varied from institution to institution, and (b) the use of campus schools for student teaching was declining. Questionnaire responses from 29 institutions yielded the following information: (a) 24 campus schools still were being used for some student teaching; (b) 19 participated in research and experimentation; (c) 14 provided for observation and participation; and (d) 11 served as locales for demonstration teaching.

From the responses to a questionnaire which was planned to determine what teachers in campus laboratory schools considered the role of such schools to be in teacher education, Venable (1960) reached the following two conclusions: (a) the student teaching function should be discontinued in campus schools, and (b) observation and participation were judged to constitute the most important functions of such a school. Rzepka (1962) noted that the purpose or function of laboratory schools, which still was not clear, continued to be a highly individual matter in terms of special needs and problems of particular teacher education institutions. He indicated that many campus schools have recast their roles from student teaching to other forms of direct experience for students as well as to research. His survey of 21 campus schools yielded the following data concerning use: (a) 21, for observation; (b) 18, for demonstration; (c) 15, for student teaching; (d) 11, for professional services; and (e) 9, for research.

The stage campus schools are in is questionable. However, there is no doubt that they are undergoing metamorphosis.

Needed Research

Much experimentation and research, especially with techniques other than the questionnaire survey, are needed in relation to all aspects of teacher education programs. Values ascribed to most of the component parts and methods used in pre-service programs rest primarily on questionable theoretical foundations. Experimental evidence is needed if a sound base for teacher education is to be established.

Continued efforts to determine experimentally the qualities and competencies required for an individual to become an effective teacher are essential to the improvement of pre-service programs. Most critical among the research needs in this field of higher education is experimentation with five-year comprehensive programs. Satisfactory evidence concerning the strengths and weaknesses of single-purpose and multipurpose institutions for achieving the goals in each of the three aspects of teacher education—general education, field of specialization, and professional education—is required to guide changes in organization and program.

Bibliography

ACKERMAN, RALPH E. *A Critical Analysis of Programs for Junior High School Teachers in Teacher Education Institutions of the United States.* Doctor's thesis. Storrs: University of Connecticut, 1960. 261 pp. Abstract: *Dissertation Abstracts* 21: 2605-2606; No. 9, 1961.

ANDREWS, L. O., and PALMER, R. R. "The Education of the Elementary School Teacher." *The Education of Teachers: New Perspectives.* Official Report of the Second Bowling Green Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1958. pp. 322-29.

COGAN, MORRIS L. "Professional Requirements in Programs for the Preparation of High School Teachers." *Journal of Teacher Education* 9: 270-79; September 1958.

FISHBURN, C. E. "Teacher Role Perception in the Secondary School." *Journal of Teacher Education* 13: 55-59; March 1962.

FOWLER, H. SEYMOUR. "Evaluation of an Institute for the Training of Elementary-School Science Resource Teachers." *Journal of Educational Research* 53: 358-59; May 1960.

GOODLAD, JOHN I., and BOOZER, HOWARD R. "Campus-Wide Committees on Teacher Education." *Educational Record* 42: 348-53; October 1961.

HAYES, ROBERT BRUCE. *A Study of Programs of Pre-Service and In-Service Education for Cooperating School Supervising Teachers Used by Selected Teacher Education Institutions.* Doctor's thesis. Lawrence: University of Kansas, 1960. 169 pp. Abstract: *Dissertation Abstracts* 21: 2611-12; No. 9, 1961.

HODENFIELD, G. K., and STINNETT, T. M. *The Education of Teachers: Conflict and Consensus.* Englewood Cliffs, N.J.: Prentice-Hall, 1961. 172 pp.

HOWELL, JOHN J., and OTHERS. *Teaching Fellows: An Experimental Program for Recruitment and Preparation of Superior Teachers.* New Haven, Conn.: Yale-Fairfield Study of Elementary Teaching, 1960. 131 pp.

HUGHES, OTTO. "The Role of the Campus Laboratory School." *Bulletin of the School of Education* (Indiana University) 35: 1-46; March 1959.

INLOW, GAIL M. "A Comparative Study of Student-Teaching Practices in Thirty-Eight Midwest Institutions." *Journal of Experimental Education* 28: 337-49; June 1960.

JONES, RODNEY M. "Off-Campus Student-Teaching Programs—Their Size and Importance." *Journal of Teacher Education* 11: 512-19; December 1960.

LEVINE, MADELINE S. "Extending Laboratory Experiences: Part II." *Journal of Teacher Education* 12: 29-35; March 1961.

LINDSEY, MARGARET, editor. *New Horizons for the Teaching Profession.* Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. Chapter 4, "Preparation of Professional Personnel," pp. 27-108.

MCAULAY, J. D. "How Much Influence Has a Co-Operating Teacher?" *Journal of Teacher Education* 11: 79-83; March 1960.

MCCREARY, ANNE PHILLIPS. "Determining Individual Needs of Elementary Teachers as a Basis for an Orientation Program." *Journal of Educational Research* 54: 20-24; September 1960.

MAYOR, JOHN R. "The Professional Education of Science and Mathematics Teachers." *Journal of Teacher Education* 13: 125; June 1962.

MEYER, KARL W. "Post-World War II Conversion of the Teachers Colleges." *Journal of Teacher Education* 11: 335-39; September 1960.

NATIONAL EDUCATION ASSOCIATION, NATIONAL COMMISSION ON TEACHER EDUCATION AND PROFESSIONAL STANDARDS. *The Education of Teachers: Curriculum Programs.* Official Report of the Kansas Conference. Washington, D.C.: the Commission, 1959. 453 pp.

NORTH CENTRAL ASSOCIATION OF COLLEGES AND SECONDARY SCHOOLS, SUBCOMMITTEE ON HUMAN RELATIONS IN THE CLASSROOM. "Teacher Education for Human Relations in the Classroom." *North Central Association Quarterly* 36: 277-304; Winter 1962.

NORTH CENTRAL ASSOCIATION OF COLLEGES AND SECONDARY SCHOOLS, SUBCOMMITTEE ON HUMAN RELATIONS IN THE CLASSROOM. "Human Relations in the Classroom: A Study of Problems and Situations Reported by 1075 Second Year Secondary School Teachers." *North Central Association Quarterly* 37: 259-79; Winter 1963.

OHLES, JOHN F. "The Laboratory School: Unresolved Problem." *Journal of Teacher Education* 12: 390-94; December 1961.

PIPER, EVELYN. *An Evaluation of the Undergraduate Elementary Teacher Education Program at the University of Oregon, Based on the Opinions of Graduates.* Doctor's thesis. Eugene: University of Oregon, 1960. 273 pp. Abstract: *Dissertation Abstracts* 21: 2209-10; No. 8, 1961.

POPHAM, W. JAMES, and GREENBERG, SUZANNE W. "Teacher Education: A Decade of Criticism." *Phi Delta Kappan* 40: 118-20; December 1958.

PRICE, ROBERT D. "The Influence of Supervising Teachers." *Journal of Teacher Education* 12: 471-75; December 1961.

RUMFORD, HERBERT PAUL. *An Experiment in Teaching Elementary School Methods via Closed Circuit Television.* Doctor's thesis. Boulder: University of Colorado, 1960. 156 pp. Abstract: *Dissertation Abstracts* 21: 2994; No. 10, 1961.

RZEPKA, LOUIS. "The Campus School: Its Search for Identity." *Journal of Teacher Education* 13: 24-29; March 1962.

SCHUMAKER, JOHN ABRAHAM. *Trends in the Education of Mathematics Teachers: A Study of the Education of Senior High School Mathematics Teachers in Selected Teacher Education Institutions in the United States in the Period 1920-1958.* Doctor's thesis. New York: New York University, 1959. 310 pp. Abstract: *Dissertation Abstracts* 20: 4044-45; No. 10, 1960.

SLOTHOWER, WILLIAM REYNOLDS. *English-Language Preparation Required of Prospective Teachers of High-School English in American Colleges and Universities.* Doctor's thesis. Urbana: University of Illinois, 1959. 161 pp. Abstract: *Dissertation Abstracts* 20: 1688; No. 5, 1959.

TAYLOR, BOB L. "Factors Influencing In-Service Teacher Education Programs." *Journal of Educational Research* 52: 336-38; May 1959.

TAYLOR, BOB L. "The In-Service Education Needs of New Teachers." *California Journal of Educational Research* 12: 221-23; November 1961.

THOMPSON, FRANKLIN JAMES. *Use of Closed Circuit Television in Teacher Education: Relationship to Professional Attitudes and Interests.* Doctor's thesis. Minneapolis: University of Minnesota, 1960. 202 pp. Abstract: *Dissertation Abstracts* 21: 819; No. 4, 1960.

TRUMP, J. LLOYD. "A Future Setting for Teacher Education." *Journal of Teacher Education* 13: 19-23; March 1962.

VENABLE, TOM C. "The Function of the Laboratory School in the Teacher Training Program." *Teachers College Journal* 32: 74-76, 88, 89; December 1960.

WIEBE, ELIAS HENRY. *The Nature of Elementary Teacher Education in Certain Liberal Arts Colleges and Universities in California.* Doctor's thesis. Los Angeles: University of Southern California, 1960. 292 pp. Abstract: *Dissertation Abstracts* 21: 1474-75; No. 6, 1960.

WILLIAMS, CHESTER S.; DEEVER, R. MERWIN; and FLYNN, EDWARD B., JR. "Professional Laboratory Experiences in Oklahoma." *Journal of Teacher Education* 11: 497-506; December 1960.

WORONOFF, ISRAEL. "Teacher Education Programs for Liberal Arts Graduates." *Journal of Teacher Education* 9: 359-62; December 1958.

YOUNG, WILLIAM M. *A Critical Analysis of Selected Research on Student Teaching in Undergraduate Teacher Training Institutions in the United States, 1929-1959.* Doctor's thesis. Philadelphia: Temple University, 1961. 170 pp. Abstract: *Dissertation Abstracts* 22: 176; No. 1, 1961.

Additional References

ABEL, FREDERICK PAUL. *Use of Closed Circuit Television in Teacher Education: Relationship to Achievement and Subject Matter Understanding.* Doctor's thesis. Minneapolis: University of Minnesota, 1960. 265 pp. Abstract: *Dissertation Abstracts* 21: 2999-3000; No. 10, 1961.

AYLESWORTH, THOMAS G., and NEWSOME, RICHARD D. "Biology for the Elementary Teacher." *American Biology Teacher* 24: 346-50; May 1962.

EGGERT, MARY MCMILLAN. *Improvement in Classroom Practices as Related to In-Service Curriculum Program.* Doctor's thesis. Salt Lake City: University of Utah, 1958. 445 pp. Abstract: *Dissertation Abstracts* 19: 1298; No. 6, 1958.

FULTON, W. R., and RUPIPER, O. J. "Observation of Teaching: Direct vs. Vicarious Experiences." *Journal of Teacher Education* 13: 157-64; June 1962.

HARRY, SHIZUKO N. "Some Trends in Teacher Education." *Educational Research Bulletin* 37: 158-60; September 1958.

HOLDEN, WILLIAM P. "The Master of Arts in Teaching at Yale, 1951-58." *Journal of Teacher Education* 10: 393-400; December 1959.

INLOW, GAIL M. "The College Supervisor of Student Teaching—A Comparative Study." *Journal of Teacher Education* 10: 211-16; June 1959.

JONES, RODNEY M. "Off-Campus Student-Teaching Programs." *Journal of Teacher Education* 12: 97-100, March; 219-22, June 1961.

KAZAMIAS, ANDREAS M. "The Education of Good Teachers and the Oberlin Master of Arts in Teaching Programs." *Journal of Teacher Education* 12: 205-208; June 1961.

NATIONAL EDUCATION ASSOCIATION, NATIONAL COMMISSION ON TEACHER EDUCATION AND PROFESSIONAL STANDARDS. *The Education of Teachers: New Perspectives.* Official Report of the Second Bowling Green Conference. Washington, D.C.: the Commission, 1958. 400 pp.

ROTH, LOIS H. "Selecting Supervising Teachers." *Journal of Teacher Education* 12: 476-81; December 1961.

SATTLER, JOHN W., and MACLOUGHLIN, ROY. "Pilot Workshop in English." *North Central Association Quarterly* 37: 194-203; Fall 1962.

CHAPTER IV

Conditions of Employment and Service in Elementary and Secondary Schools

GARFORD G. GORDON

A SURVEY of the research reported during the past five years on employment and conditions of service for teachers reveals a somewhat barren scene. The necessary routine of social bookkeeping has gone on steadily with only minor changes. Summaries of legal provisions, salary regulations, salaries paid, economic conditions, school district statistics, and measurable teacher characteristics such as age, sex, formal education, and salary received have been reported. The best and most comprehensive of these studies have been those published by the U.S. Office of Education. Many local surveys have been carried out by state education associations and other groups with special or regional interests. Most of these studies, which were directed toward legislation or toward the adoption of improved district salary schedules, generally have been of limited value to educational researchers but often of major importance to those directly involved.

When one looks for research going beyond the collection and rough classification of quantifiable facts about current conditions, the picture is rather bleak. The literature mostly consists either of statements of opinions and ideas or of descriptions of plans and panaceas—together with unevaluated and often uncritical and highly biased descriptions of innovations. Few studies which are based on clearly stated hypotheses and strict definitions of conditions have been reported.

This chapter will not discuss all of the weaknesses in the organization of educational research that contribute to this situation. However, one probable causal factor which is clearly important to the researcher has been the subject—either directly or peripherally—of some of the better studies reported since 1958. This factor is the lack of a clear definition of the role of education and of the teacher in American society. Teaching conditions, which to a great extent depend upon role expectations held by and for teachers, cannot be unequivocally evaluated or clearly analyzed when these expectations are not made explicit. This problem is revealed in varying degrees in many studies, particularly in those reported by Barr (1961), Davis (1962), Ryans (1960), and Turner and Fattu (1960).

Since their results have a bearing on the evaluation of findings in all other areas of teacher employment conditions, studies of teacher status and of role expectations are considered first in this chapter. Staff organization and utilization of staff members are treated secondly. Other topics discussed are teacher load; employment, assignment, and turnover; evaluation, retention, and dismissal; economic status and economic conditions of employment; and teacher morale and job satisfaction.

Teacher Status and Role Expectations

Most information about teacher status and role in foreign countries must be inferred from more general studies on comparative education. Research studies, as distinct from personal narratives and general descriptions of educational conditions, were lacking during the period of this REVIEW. However, there have been studies of some of the more easily determined aspects of the over-all status of public school teachers. The National Education Association, Research Division (1962d, 1963) reported findings as to the age, experience, training, and salary of teachers based on a sampling study conducted by a technique it recently had developed. Mason (1961) made a final report on the survey of new teachers conducted by the U.S. Office of Education during the period 1956-57. A follow-up study of men teachers and ex-teachers who had taken aptitude tests while in the Air Force in 1943 was reported by Thorndike and Hagen (1961). They found that ex-teachers were significantly superior to those who had remained in teaching in reading comprehension, arithmetic reasoning, and mathematics. Moreover, they were receiving an income approximately 25 percent higher than that of the men who had remained in teaching.

Although studies dealing with role expectations of teachers and with teaching were somewhat limited in coverage, they were numerous and relatively productive of implications for further study. Rugh (1961) presented a penetrating analysis of the frustrations and of the problems resulting from the public's exaggerated expectations of schools. He concluded that the public credited teachers with too much potential, set goals beyond the level of teacher competence, and indulged in wishful thinking about "what schools *should* do as against what they *can* do." Solomon (1961) analyzed the role of the teacher in public education and pointed out the contradictions inherent in the role of the professional employee. After studying the historical status of the teacher, Ulich (1958) decided that salary and quality control in the selection and in the retention of teachers would have great effect on teacher status, but that one of the determining factors in regard to the teacher's status was "the structure of the society in which he works."

Rettig and Pasamanick (1959) studied role perceptions of teachers in comparison to those of other professional workers. On the other hand, Biddle, Twyman, and Rankin (1962) compared the role expectations of teachers, education students, and noneducation students. They concluded that role distortions may be responsible for a substantial proportion of the high dropout rate during the first year of teaching. Discrepancies in role perception also were reported in studies by Rubenstein (1959) and

Fishburn (1962). Some insight into contradictions between stated and actual teacher roles was furnished by Hagstrom (1962), who found that elementary teachers spend a major part of their workday doing subprofessional tasks.

Staff Organization and Utilization

During the past five years, there has been an increase of interest in the nature and organization of teachers' work. Attempts have been made at both the elementary and the secondary levels to determine the activities required in a teaching position and to reorganize them with a view toward improving teaching conditions and the utilization of teacher time and talents. A helpful summary of research in this area was made by the National Education Association, Research Division (1959b). The best study of the use of elementary teachers' time was reported in the previously cited paper by Hagstrom (1962). He concluded that such teachers spend about 30 percent of their time at school doing subprofessional or clerical work. Stafford (1962) found that this percentage was decreased significantly when teacher aides were employed.

One change designed to reduce the noninstructional load of elementary teachers has been the introduction of provisions for duty-free noon periods. In 1959, according to a report of the National Education Association, Research Division and the American Association of School Administrators (1960), 25 of 33 urban districts surveyed reported employing lay personnel for noontime supervision of students. However, for the same year, the National Education Association, Research Division (1962a) found that only 11.9 percent of all districts over 2,500 in population had duty-free lunch periods for teachers. Most of the districts allowed for such periods simply by not providing for supervision of pupils at noontime.

Although the problems of utilization of teacher time and of the employment of teaching devices and subprofessional or clerical aides are closely interwoven with curriculum and instruction, they also have a direct bearing on the status and working conditions of teachers. Since most reports in this area were, however, descriptive or argumentative, they cannot be considered research documents. One exception was a study of opinions on the use of lay readers of English papers reported by Burke (1962).

Undoubtedly, the best publicized activity on staff utilization was conducted at the secondary level with the encouragement of the National Association of Secondary-School Principals. Most of the experiments and studies reported from this program did not constitute acceptable research. Anrig (1962) discussed many of the problems involved and concluded that there was a need to improve the standards of research being used to test such new ideas.

Teacher Load

There has been much discussion of teacher load during the period covered by this Review. Most of it has taken the form of descriptions, theories, and proposals for policies. Very little of the published material is concerned with research. Although the data presented were purely normative, those portions concerned with the determination of teacher opinion and attitudes threw some light on the relationship between load and morale or between load and job satisfaction. McLaughlin and Shea (1960) found that load was the most important source of teacher job dissatisfaction in California elementary schools.

Romine (1958) described assignment practices and their relationship to load in Colorado high schools. Using the revised Douglass formula, he found no clear trends and noted little difference between small and large high schools. Perhaps his most significant finding was that beginning teachers had about the same load as experienced teachers.

Brown and Obourn (1959) reported on assignments and teaching loads of mathematics and science teachers in three Middle Atlantic states. This was the most comprehensive survey of the teaching load of a category of teachers that could be found during the period covered by this Review. The survey was limited to use of assignment as the primary definition of load. Class size, number of classes, and other conditions affecting the load were given little attention. However, to the extent that the teacher's preparation for subject areas to which he is assigned is a factor in load, this study is quite valuable.

Other areas of concern in relation to teacher load have received less consideration than that of appropriateness of assignment. However, two factors related to load have been thoroughly surveyed, probably because they are obvious and amenable to numerical treatment. In the matter of hours of work, the National Education Association, Research Division (1963) made use of its new national sampling technique to find that the average public school teacher worked 47.3 hours a week. About 43 percent of this time was utilized in duties other than classroom instruction.

The National Education Association, which has long been interested in problems related to class size, has regularly conducted surveys, although these have been limited to elementary schools. In reports issued during the period covered by this Review, the National Education Association, Research Division (1962c) revealed that there had been some relief in middle-sized districts but that the over-all picture concerning class size had changed little since 1956.

The hypothesis that all teachers favor a class size smaller than the one they have was partially tested in a study reported by the National Education Association, Research Division (1961c). It was found that, when separately surveyed, the majority of both elementary school teachers and principals agreed that a class consisting of between 20 and 24 pupils was the optimal size.

Employment, Assignment, and Turnover

Most research on the employment of teachers reported in the last five years was oriented toward the problems of recruitment and turnover. The legal aspects of teacher employment were competently dealt with by the National Education Association, Research Division (1959a).

Green (1960) described case studies of a nationwide sample of school districts and concluded that there was an increasing use of long-range plans for recruitment. He found that larger districts tended to have more elaborate programs and to use more imaginative techniques in contrast to the small districts which followed traditional patterns. He also noted an increasing tendency to involve principals and teachers in recruitment programs. In a study of school systems around New York City, Redefer (1962a) determined that the cost of recruiting one new teacher was between $500 and $1,000. He concluded that recruitment procedures should focus on obtaining candidates qualified to fill specific needs of districts.

There was little systematic research reported on the assignment of teachers once employed. As noted previously, Romine (1958) revealed that beginning teachers were assigned the same load as were experienced teachers in Colorado high schools. Brown and Obourn (1959) found that in three Eastern states science and mathematics teachers were frequently assigned classes in subject areas other than those in which they had received their major preparation.

One of the major concerns of those individuals responsible for hiring teachers is their retention after employment. Several studies were directed toward the turnover problem among beginning teachers. Mason and Bain (1959) determined that, although teacher turnover was less than that of social workers and nurses, the average rate was 17 percent. However, this turnover rate was greatly influenced by size of district. The National Education Association, Research Division (1961b) reported a median termination rate of only 7.3 percent for districts of 100,000 or more average daily attendance. In a comprehensive survey, Mason (1961) found that 51 percent of first-year teachers did not expect to be teaching five years later, even though they were generally satisfied with their work. Kleinman (1960) discovered a significant positive correlation between the previous knowledge of the beginning teacher about conditions and his satisfaction with his work. He concluded that prospective teachers should be given detailed information about school and community conditions by districts doing recruiting.

Evaluation, Retention, and Dismissal

The major problem reported in connection with evaluation was the lack of known criterion measures having valid relationships to teacher effectiveness in producing desired changes in pupils. Moreover, many studies on teacher evaluation were reported during the period covered by this REVIEW. Davis (1962) analyzed the problems involved in securing uniformity in the assessment of teacher performance. No studies were found which dealt specifically with the processes or criteria involved in making decisions concerning retention or dismissal of staff.

In a summary of more than 75 studies on the measurement and prediction of teacher effectiveness, Barr (1961) concluded that the difficulties posed by measurement had not been overcome. However, in a major investigation, Ryans (1960) found that there were definite characteristics of teacher behavior that could be reported objectively by trained observers. However, he disregarded to a considerable extent the study of the effectiveness and of the value of the teacher behaviors that were isolated.

In a carefully conducted study, Worth (1961) raised considerable doubt concerning whether administrators were able to evaluate teachers and teaching. After completing a longitudinal study that followed the plan proposed by the AERA Committee on Teacher Effectiveness in 1952, Medley and Mitzel (1959) also concluded that their findings questioned the relevance of a considerable body of research in teacher effectiveness that had used ratings of some kind as a criterion.

The problem of teacher evaluation on the job seems to arise to a large degree because of the fact that a teacher must be effective in a given situation rather than in a generalized context. Turner and Fattu (1960) presented an insightful theoretical discussion of the research problem raised by this fact and suggested a sound proposal for attacking it. Gowan (1960) prepared a general summary of research in predicting and measuring teacher effectiveness with particular reference to the criterion problem. He also set forth a proposal for designing self-report tests that would give clues to qualities related to acceptable criteria of teaching success.

Economic Status and Economic Conditions of Employment

The major sources of data about the economic status of teachers and the economic aspects of the conditions of their employment are those of the National Education Association, Research Division and the research offices of various state education associations. Reports from state associations tend to be ephemeral or, in the librarian's terminology, "fugitive" materials. However, data on teachers' salaries and on economic conditions for the current year can be obtained for most states through a request to the appropriate state associations or from publications of the National Education Association. National reports are issued regularly by this organization and are usually available in libraries. Although the information is limited in most cases to larger school districts, the percentage of teachers omitted from their coverage is small.

The general economic status of teachers also is reviewed periodically. The latest report of the National Education Association, Research Division

(1962b) showed that there was a great variation among states and districts. Although the average salaries of teachers were found to be increasing faster than the cost of living, it was noted that less than 1 percent of all classroom teachers had a salary as high as the income estimated to be required for a family of four to live at the junior professional or executive level.

Teacher salaries are supplemented by other economic provisions of employment. Among these are retirement benefits, the legal aspects of which Punke (1962) considered in an exhaustive study. A study by the National Education Association, Research Division (1961a) showed that many districts were paying part of the premium for insurance programs covering school employees. However, Dingman (1961) found that other types of monetary benefits were rare.

Although much was written on merit rating during the period covered by this REVIEW, little had research value. The best work published was that of Stegeman (1959) who presented a competent review of the historical development of merit rating and of the differences between its application in industry and in education.

Teacher Morale and Job Satisfaction

Morale is both a result of working conditions and an indicator of the climate of human relations in which work is performed. Sometimes, the terms *morale* and *job satisfaction* are used as synonyms in the literature. There is no question that they are closely related. Perhaps it may be said that *job satisfaction* is commonly used to refer to the reactions of individuals to specific elements in their working environment, whereas *morale* often is applied to the general level of satisfaction and enthusiasm of individuals and groups.

Robinson (1959) and Robinson and Connors (1961, 1962) have summarized job satisfaction research during the period covered by this REVIEW. Some of the studies reported were concerned directly with teachers. Many others contained important implications for educational personnel administration. Blocker and Richardson (1963) gave an excellent bibliography of published and unpublished studies in this field along with a statement of needed research. They concluded that research concerning morale might provide a means for evaluation of administrative behavior—a long-standing need.

A number of studies attempted to identify factors influencing teacher morale and job satisfaction. From a study involving 24 school systems and 5,000 teachers, Redefer (1959) found that a close relationship exists between morale and the quality of education present in individual schools. Turner (1960) and Redefer (1962b) ascertained that salary was a significant morale factor. However, many other factors were also of great importance. McLaughlin and Shea (1960) found that working conditions as affected by excessive load and negative pupil attitudes were a major source of job dissatisfaction.

Some research was concerned with interactions of teachers and administrators. In a study that was an example of good research design, Willower (1960) determined that principals holding a relaxed and nondirective view of their administrative function tended to hold a higher opinion of the professional status of teachers than did those principals who appeared to be somewhat unrelaxed and directive in their orientation. Rasmussen (1962) noted that in some secondary situations discrepancies existed even when real attitudes were very similar.

Three studies indicated that discrepancies between teacher attitudes and expectancies, on the one hand, and the actual conditions in the teaching profession, on the other, may contribute significantly to morale problems in the schools. Rettig and Pasamanick (1959) concluded that the actual status of teachers is related positively to job satisfaction but that their status aspirations are related negatively to it. Day (1959) found that a very marked change in attitudes occurred in beginning elementary teachers during their first six months of service. On the other hand, Rabinowitz and Rosenbaum (1960) reported a similar change in elementary teachers who had taught three years. In both cases, the change was in the direction of increased authoritarian attitudes as measured by the *Minnesota Teacher Attitude Inventory*. All three studies raised a question about the realism of teacher training programs.

Needed Research

A major factor underlying deficiencies in research on employment practices and working conditions has been a failure to treat them as part of the general educational scene. Until their relationship to the goals of education has been better determined, little useful work can be done to evaluate them or to construct meaningful hypotheses concerning the effects of proposed changes. In particular, major research efforts are needed to develop improved definitions of teacher roles, increasingly valid criteria of teacher effectiveness, and augmented understanding of factors affecting morale. Specific research problems in some of these areas have been listed by Barr (1961); Biddle, Twyman, and Rankin (1962); Blocker and Richardson (1963); and Ryans (1960).

Bibliography

ANRIG, GREGORY R. "Promising and Perplexing Aspects of Large Group Teaching Experiments." *Bulletin of the National Association of Secondary-School Principals* 46: 253-59; January 1962.

BARR, A. S., issue editor. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness, Summary of Investigations." *Journal of Experimental Education* 30: 5-156; September 1961.

Biddle, Bruce J.; Twyman, J. Paschal; and Rankin, Earl F. "The Role of the Teacher and Occupational Choice." *School Review* 70: 191-206; Summer 1962.

Blocker, Clyde E., and Richardson, Richard C. "Twenty-Five Years of Morale Research: A Critical Review." *Journal of Educational Sociology* 36: 200-10; January 1963.

Brown, Kenneth E., and Obourn, Ellsworth S. *Qualifications and Teaching Loads of Mathematics and Science Teachers in Maryland, New Jersey, and Virginia.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 575. Washington, D.C.: Government Printing Office, 1959. 101 pp.

Burke, Virginia M. "The Lay Reader Program in Review." *Bulletin of the National Association of Secondary-School Principals* 46: 261-68; January 1962.

Davis, Hazel. "Determining Reasonable and Valid Policies and Procedures for the Continuous Evaluation of Professional Performance in Service." *Professional Imperatives: Expertness and Self-Determination.* Official Report of the Fort Collins Conference. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1962. pp. 235-59.

Day, Harry P. "Attitude Changes of Beginning Teachers After Initial Teaching Experience." *Journal of Teacher Education* 10: 326-28; September 1959.

Dingman, Erwin. "School Practices in Fringe Benefits." *American School Board Journal* 143: 7-8; August 1961.

Fishburn, C. E. "Teacher Role Perception in the Secondary School." *Journal of Teacher Education* 13: 55-59; March 1962.

Gowan, John Curtis. "Self-Report Tests in the Prediction of Teaching Effectiveness." *School Review* 68: 409-19; Winter 1960.

Green, Arthur S. "Modern Methods in Recruitment of Teachers." *American School Board Journal* 140: 13-16; May 1960.

Hagstrom, Ellis A. "The Teacher's Day." *Elementary School Journal* 62: 422-31; May 1962.

Kleinman, Lou. "A New Dimension in Teacher Selection." *Journal of Educational Sociology* 34: 24-33; September 1960.

McLaughlin, Jack W., and Shea, John T. "California Teachers' Job Dissatisfactions." *California Journal of Educational Research* 11: 216-24; November 1960.

Mason, Ward S. *The Beginning Teacher: Status and Career Orientations.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 644. Washington, D.C.: Government Printing Office, 1961. 296 pp.

Mason, Ward S., and Bain, Robert K. *Teacher Turnover in the Public Schools 1957-58.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 608. Washington, D.C.: Government Printing Office, 1959. 38 pp.

Medley, Donald M., and Mitzel, Harold E. "Some Behavioral Correlates of Teacher Effectiveness." *Journal of Educational Psychology* 50: 239-46; December 1959.

National Education Association, Research Division. *The Teacher and the Law.* Research Monograph 1959-M3. Washington, D.C.: the Association, September 1959. 90 pp. (a)

National Education Association, Research Division. *Studies of Utilization of Staff, Buildings, and Audio-Visual Aids in the Public Schools.* Research Report 1959-R17. Washington, D.C.: the Association, October 1959. 41 pp. (b)

National Education Association, Research Division. *School Districts That Pay Part of the Premiums for Certain Insurance Programs Covering School Employees.* Research Memo 1961-23. Washington, D.C.: the Association, June 1961. 13 pp. (a)

National Education Association, Research Division. *Selected Statistics of Local School Districts, 1960-61.* Research Report 1961-R20. Washington, D.C.: the Association, December 1961. 66 pp. (b)

National Education Association, Research Division. "Teachers and Principals Agree on Best Class Size." *NEA Research Bulletin* 39: 107; December 1961. (c)

National Education Association, Research Division. *Duty-Free Lunch Periods for Teachers.* Research Memo 1962-14. Washington, D.C.: the Association, April 1962. 8 pp. (a)

National Education Association, Research Division. *The Economic Status of Teachers in 1961-62.* Research Report 1962-R7. Washington, D.C.: the Association, May 1962. 54 pp. (b)

National Education Association, Research Division. *Class Size in Urban Elementary Schools, 1962.* Research Report 1962-R10. Washington, D.C.: the Association, October 1962. 25 pp. (c)

National Education Association, Research Division. "Profile of the Metropolitan Teacher." *NEA Research Bulletin* 40: 67-74; October 1962. (d)

National Education Association, Research Division. *The American Public School Teacher, 1960-61.* Research Monograph 1963-M2. Washington, D.C.: National Education Association, April 1963. 112 pp.

National Education Association, Research Division and American Association of School Administrators. *Noon-Hour Procedures in Elementary Schools.* Educational Research Service Circular No. 6. Washington, D.C.: the Association, August 1960. 17 pp.

Punke, Harold H. "Court Rulings on Teacher Retirement Benefits." *Bulletin of the National Association of Secondary-School Principals* 46: 131-79; September 1962.

Rabinowitz, William, and Rosenbaum, Ira. "Teaching Experience and Teachers' Attitudes." *Elementary School Journal* 60: 313-19; March 1960.

Rasmussen, Glen R. "Perceived Value Discrepancies of Teachers and Principals—A Threat to Creative Teaching." *Bulletin of the National Association of Secondary-School Principals* 46: 1-12; March 1962.

Redefer, Frederick L. "Factors That Affect Teacher Morale." *Nation's Schools* 63: 59-62; February 1959.

Redefer, Frederick L. "Recruiting and Selecting Teachers." *American School Board Journal* 145: 7-8; August 1962. (a)

Redefer, Frederick L. "The School Board and Teacher Morale." *American School Board Journal* 145: 5-7; July 1962. (b)

Rettig, Salomon, and Pasamanick, Benjamin. "Status and Job Satisfaction of Public School Teachers." *School and Society* 87: 113-16; March 14, 1959.

Robinson, H. Alan. "Job Satisfaction Researches of 1958." *Personnel and Guidance Journal* 37: 669-73; May 1959.

Robinson, H. Alan, and Connors, Ralph P. "Job Satisfaction Researches of 1960." *Personnel and Guidance Journal* 40: 373-77; December 1961.

Robinson, H. Alan, and Connors, Ralph P. "Job Satisfaction Researches of 1961." *Personnel and Guidance Journal* 41: 240-46; November 1962.

Romine, Stephen. "Teaching Assignments and Instructional Loads in Secondary Schools." *Bulletin of the National Association of Secondary-School Principals* 42: 55-63; November 1958.

Rubenstein, Ben O. "A Comparison Between Cultural Expectations Regarding the Rôle of the Teacher and His Actual Rôle in the Learning Process." *Educational Administration and Supervision* 45: 95-101; March 1959.

Rugh, Douglas. "The American Teacher—Victim of Role Inflation." *Journal of Teacher Education* 12: 54-56; March 1961.

Ryans, David G. *Characteristics of Teachers.* Washington, D.C.: American Council on Education, 1960. 416 pp.

Solomon, Benjamin. "A Profession Taken for Granted." *School Review* 69: 286-99; Autumn 1961.

Stafford, Curt. "Teacher Time Utilization with Teacher Aides." *Journal of Educational Research* 56: 82-88; October 1962.

Stegeman, William H. "Some Historical Comparisons of Merit Rating in Industry and Education." *California Journal of Secondary Education* 34: 205-12; April 1959.

Thorndike, Robert L., and Hagen, Elizabeth. "Men Teachers and Ex-Teachers: Some Attitudes and Traits." *Teachers College Record* 62: 306-16; January 1961.

Turner, James Leon. *Morale-Building Techniques in the Secondary Schools of Texas.* Austin: Texas Study of Secondary Education, February 1960. 20 pp.

Turner, Richard L., and Fattu, Nicholas A. "Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research." *Brlletin of the School of Education* (Indiana University) 36: 1-37; May 1960.

Ulich, Mary E. "The Historical Status of the Teacher." *Educational Forum* 22: 341-48; March 1958.

Willower, Donald J. "Leadership Styles and Leaders' Perceptions of Subordinates." *Journal of Educational Sociology* 34: 58-64; October 1960.

Worth, Walter H. "Can Administrators Rate Teachers?" *Canadian Administrator* 1: 1-6; October 1961.

CHAPTER V

Teacher Personnel Research in Higher Education

ALVIN MARKS and MARVIN N. DILLON

SINCE 1958, the majority of articles on college teaching personnel have been discursive in nature. There is still a marked lack of systematic research on problems related to teacher personnel in colleges and universities.

Need and Preparation of Teachers

The need for college teachers was spelled out clearly in an article in *Higher Education* (1961). The authors predicted a college student population in 1970 of approximately 6,000,000, of whom 4,500,000 will probably attend full time. They predicted that this enrollment would require an increase of college teachers from 232,000 during the 1959-60 academic year to 322,000 ten years later unless some new ways of teaching are discovered and accepted.

The common criterion for adequate preparation for college teaching was possession of a doctorate degree. The National Education Association, Research Division (1959) reported that only one in four new college teachers had an earned doctorate, that two in ten did not have a master's degree, that in general fewer and fewer new teachers had a doctorate, and that more and more teachers had less than a master's degree. Maul (1959) supported this point of view with evidence that the proportion of new teachers holding a doctorate had dropped from 31.4 percent during the 1953-54 academic year to 23.8 percent during the 1958-59 academic year. He predicted that colleges and universities would continue to have fewer and fewer new teachers holding top academic degrees. Stoikov (1959) disagreed with the contention that the quality of college teachers was on the downgrade because fewer new teachers possessed a doctorate. Using as a criterion the percentage of full-time teaching staff holding a doctorate rather than the percentage of only new teachers, he presented evidence that the quality of college teachers had improved. His research showed that, whereas 42.2 percent of full-time staff held a doctorate in 1953, the percentage had increased to 49 in 1956. He not only pointed out that many staff members already employed proceeded to complete the doctorate but also indicated that a selective factor probably was present among those staff members who resigned.

Alpren (1962) applied a different criterion of adequate preparation. Through surveying all 92 of the colleges and universities offering doctorate degrees in 1960, he determined the number and kinds of regularly scheduled courses that dealt primarily with the preparation of college teachers. He found that 41 of the institutions offered a total of 55 such courses and that only 23 provided one or more courses dealing with junior college teaching. (The reader should be aware of the fact that preparation of teachers for junior colleges is also conducted by institutions not offering doctorates. Consequently, these institutions were not included in Alpren's study.) He also reported in his survey that only 25 institutions provided a higher education program which consisted of four or more courses. In addition, he noted a wide diversity among these courses. Although his evidence showed a 72 percent increase since 1953 in the frequency of such courses being offered, he concluded that the broad scope of topics covered had not undergone much change.

Ratings and Evaluation

There were a few articles emerging from research on the ratings of college professors—ratings usually made by students. Pace and Stern (1958) developed an instrument called the *College Characteristics Index* (*CCI*) which was designed to describe college environments. The *CCI*, which has been used in several studies, contains a number of items asking for student perceptions of the college staff.

Thistlethwaite (1959) used the *CCI* to determine differences in college environmental press as viewed by National Merit scholars and by Certificate of Merit winners who were attending 36 different colleges and universities. Faculty characteristics viewed as desirable by the students included (a) informal and friendly relationships with students, (b) energy and enthusiasm for the subject, and (c) the relative absence of rigorous supervision of students. Thistlethwaite contrasted the ratings of faculties at schools that are productive of natural scientists to those that are productive in the arts, humanities, and social sciences. Faculties in the former as compared with the latter type of institution were viewed by the students as being more nondirective in their teaching methods but adhering more to strict curriculum requirements. Faculties from the latter group were seen as less threatening, more permissive, and more controversial than those from the former group of institutions.

Continuing with his study of talented students, Thistlethwaite (1960) asked his respondents to rate on a check list the three college instructors they thought had contributed the most to their desire to learn. The students were nearly unanimous in reporting that such instructors exhibited great enthusiasm for their subjects. More than 80 percent of the respondents mentioned that these instructors had taken a personal interest in them and in their work.

Isaacson, McKeachie, and Milholland (1963) described a provocative research effort in which a group of college professors did peer ratings of each other and in which their students also rated them. Using correlation techniques, the authors found that on the whole the results were inconclusive. However, enough consistency was found among the variables to indicate that a college teacher who has an agreeable nature, gives evidence of being stable, shows enthusiasm, and reflects high cultural attainment has a substantial likelihood of being regarded as an outstanding teacher by both his students and his colleagues.

In surveying all the available Ph.D. graduates from Peabody College since 1919, Hedges (1962) asked them to respond to the question, "What were the major attributes of the professor who exerted the most positive influence on you in college?" Since 41.2 percent of the replies related in some way to the professor-student relationship, there was the suggestion that respondents believed that they were influenced more by personal-social factors than by any specific teaching technique.

Veldman, Peck, and McGuire (1961) proposed an interesting way to investigate the personal construct system of professors. They utilized a sample of educational psychologists who had been asked to rate their students on a scale of over-all mental health. The ratings of the judges then were submitted to factor analysis to see whether individual construct systems would emerge and, if so, whether they would indicate significant differences. Individual construct systems were found, and some significant differences were noted. The authors, who pointed out that the general logic of such a technique could be applied to any kind of judging procedure, proposed calling it the "J technique."

Several college professors, who were found to be reluctant about being rated on subjectively judged factors, pointed out that such ratings were biased by the raters' own attitudes. Weaver (1960) and Chansky (1958) found some support for this position.

Weaver wished to examine the relationship between the grade which a student expected to receive and his rating of the instructor. He also investigated whether student criticism would be directed toward the instructor's personality, toward his teaching skills, or toward both. He found that students' ratings of their instructors were biased in the directions of the grades that the students expected to receive. More student bias was directed toward the teaching skills of the instructors than toward their personality characteristics.

Chansky used the *Minnesota Teacher Attitude Inventory* to place students into an authoritarian and into a democratic group. He then asked these students to rate their instructors to determine whether their own frames of reference would bias their perceptions. He found that democratic raters were inclined to give democratic ratings, and that authoritarian raters were apt to perceive their instructors as authoritarian.

A limited number of new studies appeared in which an attempt was made to evaluate various methods of increasing the effectiveness of college professors. The majority of such studies continued to reveal little difference in achievement that could be attributed to any particular method. The investigations by Churchill and John (1958) and by Marr and others (1960) were examples of such studies. Additional information concerning aspects of teaching effectiveness may be found in the volume edited by Gage (1963).

The U.S. Department of Health, Education, and Welfare, Office of Education (1960) advanced the hypothesis that class size is not the important variable in teaching effectiveness in higher education but that the quality of the teaching and learning is. This hypothesis was supported by 12 papers which were reviewed in the publication. Similarly, the publication posited that no one method could be demonstrated to produce more or better learning than any other. This hypothesis was sustained in terms of the findings in eight papers. Interestingly enough, despite the latter hypothesis, the publication contained a further hypothesis that approaches to teaching that are problem-oriented improve learning. Positive evidence for this hypothesis was explicit in the reviews of 12 studies.

Salaries and Other Benefits

The National Education Association, Research Division (1962) reported that the median salary paid by universities and colleges in 1961-62 to all teachers for the nine-month school year was $7,486. The median for men exceeded that for women by $1,289. One-fourth of all the men received $9,390 or more, and one-fourth received $6,451 or less.

Bokelman and D'Amico (1963) compared mean salary by academic rank between public and private institutions of higher learning for the 1961-62 academic year with the following results:

Rank	Public	Private
Full Professor	$10,200	$10,200
Associate Professor	8,280	7,735
Assistant Professor	6,960	6,495
Instructor	5,770	5,445

In another study, Bokelman and D'Amico (1961) reported average salaries of selected faculty groups which showed that higher salaries were received by the more specialized and by the less numerous groups. In general, the highest salaries appeared for the medically related faculties.

The American Association of University Professors, Committee Z (1962) presented the results of their latest biennial salary survey of a group of 41 selected colleges and universities. The authors broke down their data by geographical regions and by types of universities. The survey, which was first conducted in 1939, has been repeated faithfully every two years since the academic year 1948-49. No detailed attempt will be made here to synthesize and to summarize the data. However, the over-all salary reports did not vary greatly from those of Bokelman and D'Amico. Since most salary figures quoted included institutional contributions to retirement funds, the reader needs to keep this fact in mind in attempting to make any direct comparisons.

The study showed that, with respect to dollar amounts and over-all averages, salary advances for this biennium compared well with those of the preceding biennium. However, when consideration was given to the rate of progress being made toward restoring to instructional staffs their relative economic status of pre-World War II days, it was apparent that salaries at these selected institutions were falling behind those of the preceding two-year period. The study further showed that, if instructional salaries were to regain their competitive position by the 1967-68 academic year, there would have to be an average increment of not less than 7 percent per year, cumulatively measured. The 41 selected colleges and universities had achieved a cumulative rate of only 5 percent in this biennium.

The data made it quite clear that there had been a leveling between ranks. Possibly because of competition with others who desired young college graduates, the mean base salary for instructors had risen more rapidly than the mean base salary for other ranks. Not only the spread in compensation from the lowest to the highest ranks had decreased but also the percentage of professors receiving top salaries had declined. During the 1939-40 academic year, there were 47 professors out of 2,511 receiving $10,000 or more. With adjustment for cost of living, it would take $21,500 in 1962 to equal the 1940 figure. If the same proportion were to receive this high salary, there would need to be about 116 professors at this level; there were only 13.

Heim and Baumol (1962) presented a digest of information gleaned from the annual self-grading compensation survey for the 1961-62 academic year which was conducted by the American Association of University Professors. The survey was quite complete. Suggestions were made regarding ways in which a professor might be able to maximize his fringe benefits.

To summarize, the fringe benefits available to college professors were not high when compared with those of American business executives. The median institution contributed about 8 percent of salary in the form of countable benefits. Low benefit-salary ratios of 6 percent or less of average salary were paid by 20 percent of the institutions. High benefit-salary ratios of 10 percent or more were paid by 30 percent of the institutions.

In general, public institutions tended to lag behind private institutions. Ludlow (1959) reported that new college teachers were attracted most to individual colleges and universities by the time made available for and by the encouragement given to research and writing. Other important criteria were type and size of institution, geographic location, library facilities, salary, and opportunity for promotion. Lowest in attractive value were academic rank, sabbatical leave privileges, and retirement benefits.

Orlans (1962) mentioned that 68 percent of federal money went to 25 major universities and that, although 186 private liberal arts colleges participated in federally financed programs of research, their share of such funds was only 1.1 percent. This definitive report is a rich source of information about the effect of federal programs on teacher personnel problems.

An excellent summary of much of the pertinent data from the U.S. Office of Education's 1961-62 survey of faculty salaries was given in *School and Society* (1962).

New Data Sources

There are three sources of data that may be overlooked because of their recency and because of their somewhat restricted distribution. The U.S. Office of Health, Education, and Welfare undertook two ambitious publications: *The Reporter* and *New Dimensions in Higher Education. The Reporter*, which through April 1962 has appeared five times, consists of well-annotated bibliographies of all investigations reported to a clearinghouse for studies on higher education. Similarly, only 10 issues of *New Dimensions in Higher Education* have been published. Each number was intended to survey in summary form a specific area of concern in higher education. Studies supplied to the clearinghouse provided the basic background data.

The American Educational Research Association, a department of the National Education Association, published a collection of research studies on teaching in the *Handbook of Research on Teaching*, edited by Gage (1963). Chapter 7, "Rating Methods in Research on Teaching," by Remmers, and Chapter 23, "Research on Teaching at the College Level," by McKeachie were particularly relevant to the subject under consideration.

Bibliography

ALPREN, MORTON. "A Survey of Courses and Programs for the Preparation of College Teachers." *Journal of Teacher Education* 13: 272-78; September 1962.

AMERICAN ASSOCIATION OF UNIVERSITY PROFESSORS, SUBCOMMITTEE ON THE BIENNIAL SALARY SURVEY OF COMMITTEE Z ON THE ECONOMIC STATUS OF THE PROFESSION. "Instructional Salaries in 41 Selected Colleges and Universities for the Academic Year 1961-62." *AAUP Bulletin* 48: 19-42; March 1962.

BOKELMAN, W. ROBERT, and D'AMICO, LOUIS A. "A Comparison of 1959-60 Average Salaries of Selected Faculty Groups." *College and University* 36: 311-17; Spring 1961.

BOKELMAN, W. ROBERT, and D'AMICO, LOUIS A. "Mean, Maximum, and Minimum Salaries of Professors in Undergraduate Four-Year Colleges and Universities: 1961-62." *AAUP Bulletin* 49: 63-65; March 1963.

CHANSKY, NORMAN M. "The Attitudes Students Assign to Their Teacher." *Journal of Educational Psychology* 49: 13-16; February 1958.

CHURCHILL, RUTH, and JOHN, PAULA. "Conservation of Teaching Time Through the Use of Lecture Classes and Student Assistants." *Journal of Educational Psychology* 49: 324-27; December 1958.

·GAGE, N. L., editor. *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. 1218 pp.

Hedges, William D. "The Effective Teacher at the Doctoral Level in Education." *Peabody Journal of Education* 39: 351-54; May 1962.

Heim, Peggy, and Baumol, William J. "On the Fringe: The State of Retirement Contributions and Other Nonpecuniary Faculty Benefits." *AAUP Bulletin* 48: 346-58; December 1962.

Higher Education. "Ten-Year Objectives in Higher Education Staffing 1960-61 Through 1969-70." *Higher Education* 17: 3-14; February-March 1961.

Isaacson, Robert L.; McKeachie, Wilbert J.; and Milholland, John E. "Correlation of Teacher Personality Variables and Student Ratings." *Journal of Educational Psychology* 54: 110-17; April 1963.

Ludlow, H. Glenn. "What Attracts the New College Teacher?" *School and Society* 87: 323; June 20, 1959.

Marr, John N., and others. "The Contribution of the Lecture to College Teaching." *Journal of Educational Psychology* 51: 277-84; October 1960.

Maul, Ray C. "Will New College Teachers Be Adequately Prepared?" *Educational Record* 40: 326-29; October 1959.

National Education Association, Research Division. "Are New College Teachers Prepared?" *NEA Research Bulletin* 37: 87-92; October 1959.

National Education Association, Research Division. *Salaries Paid and Salary Practices in Universities, Colleges, and Junior Colleges, 1961-62.* Research Report 1962-R2. Washington, D.C.: the Association, February 1962. 59 pp.

Orlans, Harold. *The Effects of Federal Programs on Higher Education.* Washington, D.C.: Brookings Institution, 1962. 361 pp.

Pace, C. Robert, and Stern, George G. "An Approach to the Measurement of Psychological Characteristics of College Environments." *Journal of Educational Psychology* 49: 269-77; October 1958.

School and Society. "Faculty Salaries and Student Costs." *School and Society* 90: 334; October 20, 1962.

Stoikov, Vladimir. "Has the Quality of the College Teacher Declined?" *Journal of Higher Education* 30: 455-58; November 1959.

Thistlethwaite, Donald L. "College Press and Student Achievement." *Journal of Educational Psychology* 50: 183-91; October 1959.

Thistlethwaite, Donald L. "College Press and Changes in Study Plans of Talented Students." *Journal of Educational Psychology* 51: 222-34; August 1960.

U.S. Department of Health, Education, and Welfare, Office of Education. "Effectiveness in Teaching." *New Dimensions in Higher Education.* No. 2. Washington, D.C.: Government Printing Office, 1960. 28 pp.

Veldman, Donald J.; Peck, Robert; and McGuire, Carson. "Measuring the Value Systems of Education Professors." *Journal of Educational Psychology* 52: 330-34; December 1961.

Weaver, Carl H. "Instructor Rating by College Students." *Journal of Educational Psychology* 51: 21-25; February 1960.

CHAPTER VI

Economic, Legal, and Social Status of Teachers

HAZEL DAVIS, MARTHA L. WARE, FRIEDA S. SHAPIRO, ELEANOR DONALD, and GERTRUDE N. STIEBER *

THREE sharply contrasting types of research are reported in this chapter. Studies of the economic status of the teacher include primarily statistical surveys of practice with a limited amount of theoretical discussion. The basic information came largely from questionnaires and public records. Research on legal status was done by documentary analysis; questionnaires seldom were used. Many of the studies of economic and legal status of teachers were the work of state and national professional organizations.

This chapter also reviews research studies concerning social status as well as the related topics of teachers' attitudes, opinions, and morale. In contrast to published papers regarding legal and economic status, those pertaining to social status and related problems represent the work primarily of psychologists and sociologists who collected data chiefly through interviews, observations, and questionnaires.

There are few studies considered under more than one topic. Many significant studies were omitted because of space limitations.

Economic Status

Salaries

Research on the economic status of teachers continued to put major emphasis on salary gains and losses in comparison with other occupational groups. For example, in an annual study, the National Education Association, Research Division (1962b) presented data on trends in teachers' salaries and compared teachers' salaries with earnings of other professional occupations. Presenting data for the 1961-62 school year, the study revealed that, since 1950, salaries of the instructional staff, which includes classroom teachers, principals, librarians, and building supervisors, gradually have increased. By 1961, the average salary of the instructional staff was 14 percent above the average earnings of all wage-and-salary workers. Moreover, for the first time since 1940, salaries of the instructional staff had caught up with the earnings of the average employee in manufacturing.

---

* The material presented in this chapter represents a condensation of a longer review of research which, for reasons of space limitations, the editor found necessary to shorten. He assumes full responsibility for deletion of references or for any change in emphasis or interpretations that may have been introduced.

This comparison was between two groups, one of which consisted of professionally trained workers, most of whom were college graduates, many with two or more degrees. The other was composed of wage earners in manufacturing, government, and all industries, many of whom were unskilled and semiskilled laborers and clerical or service workers with no professional training or responsibilities.

The tendency to make this type of comparison has been due to the lack of continuing sources of information about professional salaries in general. A series on salaries for selected professional occupations has been released annually by the U.S. Department of Labor, Bureau of Labor Statistics, since 1961. This series covers salaries of 36 classifications of accountants, auditors, attorneys, chemists, and engineers working in private industry. The latest issue (U.S. Department of Labor, Bureau of Labor Statistics, 1962) showed that the mean salary of Class I auditors was less than the estimated average annual salary of $5,719 paid to classroom teachers in the 1960-61 academic year (National Education Association, Research Division, 1961a).

The National Education Association, Research Division (1962c) estimated that the average salary paid classroom teachers at all levels of preparation and experience during the 1962-63 school year was $5,735. It further was estimated that 23.5 percent of all classroom teachers in 1962-63 were paid less than $4,500, but only 10.5 percent received $7,500 or more. The average salary for the entire instructional staff during the 1962-63 school year was estimated at $5,940.

In a 1960 study of the median salaries of chemists, Roethel (1961) found that in each of six types of positions held by men chemists with bachelor's degrees, master's degrees, and doctor's degrees, respectively, men chemists who were teaching chemistry were earning considerably less than men chemists who were employed in research, development and design, production and inspection, marketing and sales, or administration and management.

Teachers' salaries, however, have been rising. Although they are not keeping up relatively with increases in earnings of many other professional groups, they are rising more rapidly than the salaries of employees of other branches of state and local governments. From figures reported by the U.S. Department of Commerce, Office of Business Economics (1962), salaries of nonschool employees of state and local governments increased 54.5 percent from 1951 through 1961, compared to an increase of 70 percent for employees in public education.

Salaries of Women

The U.S. Department of Labor, Women's Bureau (1963) prepared a pamphlet on the status of women workers. The publication stated that almost 3 million of the more than 22 million women employed in 1962 were performing professional and technical work in April 1962. This number was 58 percent more than that in 1950 and 87 percent more than the estimate for 1940. However, the proportion of women professional workers in the total women labor force was the same in 1962 as it was in 1940. The report also stated that the number of women professional workers was exceeded only by that of clerical workers, service workers, and operatives. The 1.3 million women teachers—excluding college instructors—at work in April 1962 accounted for 43 percent of all professional women workers.

In 1961, women held 36 percent of all professional positions in the United States. The U.S. Department of Labor, Women's Bureau (1962) reported on the degree of salary discrimination against women as compared with men in the same positions and with the same training. The median salary for professional workers in 1960 was $4,384 for women and $6,848 for men.

In view of this difference between professional men's and professional women's salaries, it is not surprising that, since women accounted for approximately 69 percent of all teachers, salaries of teachers were lower than earnings for other professionals (National Education Association, Research Division, 1963a). However, the differences between teachers' salaries and the earnings of other professional workers were not due entirely to sex.

Income from Additional Employment

The National Education Association, Research Division (1963a) reported that 27.3 percent of the teachers held additional jobs during the school year. Larger proportions of men than of women reported either a summer or a supplementary school-year job or both—72.5 percent for men compared with 15.5 percent for women.

Selected Research Studies

The National Education Association, Research Division and most state education associations conducted extensive studies regarding salaries of teachers and of school administrators as well as surveys concerning salary scheduling.

The National Education Association, Research Division (1962f) reported that at the beginning of the 1962-63 school year public school teachers in 32 states were employed under the provisions of minimum salary laws. Average salaries of teachers in each state are reported annually by the National Education Association, Research Division, a typical example of which was a 1962 publication (National Education Association, Research Division, 1962c).

In its most recent biennial salary study, the National Education Association, Research Division (1963d) reported on the average salary for all classroom teachers in terms of size of school system and in terms of other classifications. Average national salaries also were shown for certain other personnel, including principals, counselors, and superintendents. Salaries of other personnel groups were reported for large public school systems only.

Two recent research studies on salary schedules of classroom teachers and of administrative employees were those of the National Education Association, Research Division (1962a, 1963b). For the 1962-63 school year, these two investigations represented a first attempt to obtain salary data on the basis of enrollment of school systems. For both studies, reports were compiled from data for systems with enrollments of 6,000 or more.

In a research study for the RAND Corporation, Kershaw and McKean (1962) discussed the relationship between the typical salary schedule and teacher shortages and other problems affecting education. They proposed that salary schedules include supplements for course credits in certain subject matter fields as determined by local boards of education. These additional amounts would enable the schools to attract well-trained persons to fields with severe shortages.

Merit pay as a feature in teachers' salary schedules continued to be a topic of controversy and study. One report by the National Education Association, Research Division (1961e) dealt with the status of merit pay plans and with reasons why many have been discontinued.

Legal Status

To determine the legal status of public school teachers, one finds it necessary to analyze constitutional and statutory provisions, school board rules and regulations, and judicial decisions. The most comprehensive study of the legal status of teachers, which included consideration of both statutory provisions and judicial decisions, was *The Teacher and the Law* (National Education Association, Research Division, 1959). Topics included were admission to the profession, employment conditions, relationships with pupils, retirement, and collective bargaining.

The legal principles that govern the status of teachers as evolved through judicial decisions received extensive treatment in two recently revised school law textbooks—one by Hamilton and Mort (1959) and the other by Remmlein (1962). In addition to case law material, Remmlein included examples of statutes to illustrate the legislative approach to particular topics.

The National Education Association, Research Division published yearly digests of court decisions in which teachers were litigants. The most recent report was for the year 1962 (National Education Association, Research Division, 1963c). The latest annual review of significant court cases involving teachers was prepared by Garber (1962). Compilation each year of important state school legislation by the National Education Association, Research Division furnished up-to-date developments on enactments of interest and concern to teachers. Several of the significant outcomes in school legislation for the period from January 1 to August 1, 1962, were cited in a research report of the National Education Association, Research Division (1962d).

Remmlein and Ware (1959) used an opinion poll to evaluate the laws concerning protection of teacher personnel. Staffs of state education departments, officers of state school board associations, state education associations, local school staffs, and college professors rated the laws highly.

Additional Reference: Gauerke (1959).

Certification, Tenure, and Autonomy

Armstrong and Stinnett (1962) issued the latest biennial manual on certification requirements for school personnel as fixed by statutes and regulations in all the states. This report showed upgraded requirements for preparation; in 43 states and in the District of Columbia, a bachelor's degree was mandatory for the lowest regular initial certificate for elementary school teachers.

Ligtenberg (1958) surveyed tenure court decisions and arrived at two conclusions relative to the past three decades: (a) The courts have held that the purpose of tenure legislation is to improve the school system by protecting competent teachers from dismissal for political, partisan, capricious, or purely personal or malicious reasons. (b) The courts have insisted on strict adherence to the prescribed procedures in accordance with their interpretation of the intent of the tenure laws as providing an orderly process of dismissal and in accordance with their position that the substances of these laws lie in the right of the teacher to insist on strict observance of the procedures.

Geis and Huston (1962) expressed the view that despite the existence of tenure laws the right of teachers to continue in their jobs as long as their classroom performance is satisfactory has never been adequately defined or established. Their study examined the present state and the emerging trends in the dismissal of tenure teachers for reasons related to personal extracurricular acts or to alleged classroom misconduct.

Statutory summaries of state tenure and contract laws have been included in the School Law Summaries that were prepared by the National Education Association, Research Division (1962g).

The responsibility of the teaching profession for assuming autonomy for the advancement of professional standards was a major concern of the task force of the National Commission on Teacher Education and Professional Standards in their two-year project, *New Horizons for the Teaching Profession* (Lindsey, 1961). The report recommended that national, state, and local education associations create professional standards commissions to establish and to enforce standards for the admission and retention of qualified persons in teaching and for the separation of the incompetent and unethical. The belief was expressed that this aim would be furthered by the legal sanction of "professional standards acts."

The goal of the organized teaching profession for self-discipline and self-determination in matters of teaching competence and professional conduct was advanced when Kentucky pioneered in adopting a professional standards act (National Education Association, Research Division, 1962d). A commission created by this law was given responsibility to develop and to administer professional practices criteria through the teaching profession.

Another step in the realization of the above-mentioned goal was the unique feature in the California tenure law which recognizes (a) expert testimony in dismissal actions before the courts on questions of competency, professional conduct, and fitness to teach and (b) the professional standards commission of a state-wide professional teachers organization as a source of such expert testimony. Williamson (1960) reported not only that the procedure of expert panel investigation and testimony proved sound and workable but also that its acceptance by a court in a teacher dismissal action set an important precedent.

Civil Rights

Bolmeier (1960) discussed the legal scope of teachers' freedoms as revealed by court decisions in four areas: (a) subversive political activities, (b) participation in political affairs, (c) membership in teachers unions along with the right to strike, and (d) outside activities of teachers. He concluded that teachers were restricted more than most citizens in the exercise of freedoms guaranteed by constitutional provisions and that one should become reconciled to these legal limitations before entering the teaching profession.

The search of the National Education Association, Research Division (1961c) into restrictive and protective legal provisions relating to political activities of public school teachers brought to light few specific statutory restrictions that impede teachers from participating in political affairs at the local, state, or national levels. It was stated that, if restrictions do exist, they may be based on written or unwritten school board policies or, perhaps, on the pressure of public opinion. Analysis of judicial opinions indicated that courts are not likely to uphold dismissal, suspension, or nonappointment of politically active teachers so long as the activity does not interfere with the teacher's efficiency, takes place outside school hours, and does not violate limiting state statutes.

Bryson (1961) studied the legality of loyalty oath requirements of public school teachers. He found that prescribed methods for insuring loyalty qualifications among teachers have been considered reasonable by the courts and that the loyalty laws have not been judicially regarded as violating constitutional guarantees of freedom of speech, belief, press, and assembly.

Television, Radio, and Programed Instruction

Legal problems relating to residual rights, royalties, and compensation of classroom teachers for use and reuse of instructional materials in radio and television broadcasting and in programed instruction were considered by the National Education Association, Commission on Professional Rights and Responsibilities and the Division of Audiovisual Instructional Service. The dimensions of the problems as well as approaches and guidelines to a fair policy for teachers appeared in a publication of proceedings of a conference (National Education Association, Division of Audiovisual Instructional Service, 1962).

Teacher-Board Relations

Formal relations between teachers organizations and boards of education were the subject of increasing interest during the period reviewed. Professional negotiation and collective bargaining were vying for teacher support. The special memo on collective bargaining prepared by the National Education Association, Research Division (1958) analyzed state statutes and court decisions involving teachers and other public employees. Much of this information was brought up to date in a later publication on professional negotiation (National Education Association, Research Division, 1962e). Besides setting forth the procedures for professional negotiation and its differences from collective bargaining, the latter publication included an appendix containing a state-by-state digest of statutes, court decisions, and opinions of attorneys general on the issue involved.

Teacher-Pupil Relations

Legal liability of school districts and school personnel for pupil injuries as well as for measures to prevent accidents in all areas of school activities received attention in a comprehensive publication prepared by the National Education Association, Research Division (1963e) for the National Commission on Safety Education. It explains the law of negligence as it relates to school personnel, summarizes statutes and court decisions concerned with school district liability, and emphasizes the responsibility of the district, the administrator, and the classroom teacher for pupil safety.

In stressing safety as an instrument in preventing liability for negligence, Kigin (1963) directed his study mainly to the school shop teacher. He stated that passage of mandatory legislation to indemnify the teacher against financial loss in pupil-injury cases appeared to be the most equitable solution yet developed. He suggested a plan for pupil compensation on a state-wide basis, akin to workmen's compensation, with the state education department acting as a self-insurer.

Although the release of pupil records to parents and others has long been an educational issue, it has begun to emerge as a legal issue as well. Adams (1962) discussed the legal status of the counselor who often is faced with the problem of the release of records. To date, the law does not seem to offer teachers definite answers in this increasingly important field. The most comprehensive study, which included many court citations, was done by Bakken (1961). Further research is needed. There is a tendency for writers to confuse the legal consequences to the teacher for voluntarily releasing records with the question of who has the power to force the teacher to release records that he may not wish to relinquish.

Retirement

Two studies treated the subject of retirement extensively. A study by the National Education Association, Research Division (1961b) included data on 64 state and local systems to which teachers belong. It reported that $7.3 billion was invested in these systems and that there were more than 1.85 million active members of the systems, nearly 1 million of whom also were covered by social security. Krislov (1962) surveyed public retirement systems whose members also were covered by the federal Old-Age, Survivors, and Disability Insurance (OASDI). He found that those retiring after January 1961 received substantially higher combined benefits than those who retired before social security coverage was added. He also found that coverage through OASDI resulted in improved disability and survivor protection and that retirement systems have improved substantially since the addition of OASDI coverage.

The National Education Association, Research Division (1960) presented a digest of 54 reported retirement court decisions during the 10 years between 1950-60 in a publication similar to others issued in 1940-50. The cases since 1950 fell into two general categories: (a) disagreements with administrative determinations and (b) challenges to legislation repealing or modifying existing provisions in retirement statutes.

A current topic of interest to teachers—the tax-sheltered annuity—was a subject included in the *Proceedings of the Fortieth Annual Meeting* of the National Education Association, National Council on Teacher Retirement (1962). It was reported that some teachers in at least 20 states were participating in tax-sheltered annuity programs. The tax-sheltered annuity was the subject of three issues of the Retirement Income Series

(National Education Association, Research Division, 1961d), a series on retirement and social security which is published irregularly. Teachers were advised to seek competent advice before embarking on a tax-sheltered program.

The transferability aspects of federal, state, and local government retirement laws and policies were the subject of a report of the Advisory Commission on Intergovernmental Relations (1963). Although most of the public employee retirement systems reviewed had vesting provisions, the majority of the 60 systems with such features required 15 or more years of service before vesting. It was noted that these lengthy service requirements sometimes inhibit public employee mobility when it is in the public interest. The report showed that, of the 19 states providing for intrastate employee transferability, most were between state teachers and state employees systems and between state teachers and local teachers systems. The right to purchase out-of-state public service credit was available in 28 state teachers retirement systems. The Commission thought that provision should be made for public employees to change jobs without suffering major loss of retirement benefits. Intrastate reciprocal retirement laws and early vesting after service of no more than five years in the retirement system were among the recommendations of the Commission.

Social Status

Social status is a rather nebulous characteristic of an individual or group; it may be determined by one or by a combination of indicators. Some of these indicators for teachers are economic level, educational level, occupation of father, and images and stereotypes of teachers and teaching as an occupation. Some of these indicators tend to lower teachers' social status; others tend to raise it.

Various aspects of economic status of teachers—salary, employment status of spouse, extra-teaching jobs held—were previously discussed. Clearly, the relatively modest economic status of teachers tends to lower their over-all social status.

Educational Level

In a synthesis of research on social mobility, Lipset and Bendix (1959) observed that positions of leadership and social responsibility in any society usually are ranked at the top in a hierarchy of prestige, with positions requiring long training and superior intelligence ranked just below. Many teachers have improved their status through educational attainment.

Mason (1961) found a relationship between the education achieved by the beginning teacher and his social origin: the larger the community of origin, the better educated the teacher. He also found that in general the children of white collar workers were slightly better educated than were the children of blue collar workers.

Occupation of Father

A traditional determinant of social status is father's occupation. Mason (1961) reported that fathers of beginning teachers, to a slightly greater extent than those of all teachers, were white collar workers. The proportion of beginning teachers who were children of white collar workers was smaller than that of the average of all professionals. The social background of beginning teachers varied in different regions as well as in different types of school districts.

The National Education Association, Research Division (1963a), which reported on the occupational background of teachers' fathers, showed that teachers entering the profession in five time periods—before 1905, 1905-14, 1915-24, 1925-34, and after 1934—tended to reflect, in paternal status, the changes in occupational distribution of the male labor force through the 50-year period. There was, however, a larger proportion of teachers from the managerial and professional class and a smaller proportion from skilled and unskilled labor than in the labor force as a whole.

Lipset and Bendix (1959) indicated that studies of social mobility that do no more than compare the father's occupation with the occupation of the respondents miss significant aspects. Carlson (1961) maintained that the father's occupation was an invalid indicator of a teacher's social status. Reporting on a sample in the San Francisco area, he found that teachers had origins at all levels of the social class continuum. The proportion of teachers who were in the upper half of the social class continuum exceeded by far the proportion from the lower half. Male secondary school teachers were lowest in terms of social class origin, and female elementary school teachers were highest.

Images and Stereotypes of Teachers and Teaching

One aspect of the social status of teachers is reflected in their images as seen by themselves, by those involved in a choice of career, and by the public. Furness (1960) and Belok and Dowling (1961) surveyed novels and other literary forms for characterizations or stereotypes of teachers. Furness reported distorted stereotypes of high school teachers and observed that these distorted stereotypes represented a dangerous cultural lag. Although fiction often is ahead of fact, it lagged behind reality in the literature surveyed. Although Belok and Dowling found less evidence of definite stereotypes of college professors, they noted a disproportionate number of undesirable characterizations and concluded that college teaching was in the strange position of being highly honored yet constantly ridiculed.

Kuhlen and Dipboye (1959) attempted to discover ways in which prospective teachers differed from prospective members of certain other occupations. They found that public school teaching was perceived as offering very low potential for substantial economic gain; that those selecting teaching had higher needs for nurturance, deference, self-abasement, order, and affiliation than those choosing other professions; but that they exhibited lower needs for achievement, autonomy, and change than other professionally directed groups.

In his analysis of the professional role perceptions of teachers who expressed evaluation of their own activities, Drabick (1959) found that elementary school teachers generally did least and that special teachers did most to create an impression of an educational profession. He proposed that until teachers engaged in more out-of-school activities which they expressed as being professional, they would have little claim to be regarded as members of a profession as are doctors, lawyers, and ministers.

In what might be viewed as an indirect approach to the determination of the stereotype of the teacher held by students, Devita and Kaczkowski (1959) polled a random sample of high school seniors in the Chicago area to determine their reasons for and against selecting teaching as a career. In rank order of frequency of citation, the negative reasons were classified as follows: (a) interest in other careers, (b) expectation of low salary, (c) lack of appeal which teaching held, (d) perceived lack in their own ability, and (e) uninteresting nature of the work involved in teaching. Reasons for choosing teaching were as follows: (a) pleasure in working with children, (b) belief that teaching was enjoyable work, (c) desire to teach a particular grade or subject, (d) enjoyment in helping others, and (e) perception of the need for teachers.

O'Dowd and Beardslee (1961) studied the concept of the male school teacher as seen by college undergraduates. They found that teaching as a professional role for men is highly regarded by college undergraduates; men rated it higher than did women. Many features of the personality, style of life, and patterns of social participation of the school teacher appeared to be appealing to the young men, although they also recognized the unfavorable economic factors.

From a study of the behavior of public school teachers in community life, Buck (1960) reported that participation was a response to community expectancy and that its frequency was to a large extent related to the varying and sometimes contradictory pressures from the community to take part in its activities. He found little consensus on the definition of the teacher's role as it related to students. Moreover, teachers who achieved a workable definition of role attained their objectives despite varying and contradictory expectations.

In one community, Chilcott (1961) found a rather formless teacher stereotype. He conjectured that the traditional teacher stereotype of 25 years ago has broken down and has not been replaced yet. He found some evidence that the new stereotype may be that of an attractive and youthful married woman who commands considerable community respect.

Groff (1962) reviewed research on social status of teachers since 1925. He found that it was difficult to fit teachers into widely accepted social status scales and that teachers might be classified best as "middle class professionals."

Additional References: Charters (1963); Wilson and Goethals (1960).

Morale, Opinions, and Attitudes of Teachers

Studies of morale, opinions, and attitudes of teachers were of three basic types: (a) those which were purely descriptive of the level of morale or of teachers' attitudes toward a particular item or several items, (b) those which attempted to correlate morale with various casual factors, and (c) those which focused primarily on various methods of measuring level of morale. Since many studies involved two or more types of the emphases just described, no effort has been made to classify them.

Three papers appeared that presented overviews of survey-type research concerning teacher morale. Summarizing studies made in 20 school surveys, Harap (1959) found that the most common causes of poor morale were inadequate salary, large classes, poor administration, lack of a daily period of relaxation, unsatisfactory plant, and lack of teaching materials and equipment.

Redefer (1959) reported a summary of opinionnaire studies in which 5,000 teachers in 24 systems participated. He concluded that action to improve teacher morale should be concentrated on (a) board of education and administrative relations, (b) personnel practices and policies, (c) provision for adequate school equipment and supplies, and (d) educational leadership. Redefer (1963) reached a similar conclusion in a later review of more than a score of doctoral studies.

Kuhlen (1959) described career patterns in public school teaching and sought to identify satisfactions and frustrations. He attempted to show the relation between the psychological needs of individual teachers and their perceptions of the degree to which those needs might be satisfied in the teaching profession, as well as the relation between the psychological needs of teachers and their general satisfaction with their careers and with their occupations. The importance of the relationship between the psychological needs of prospective teachers and a realistic perception of the extent to which these needs might be met was rejected in a study by Day

(1959). He found that college students preparing for teaching exhibited unrealistic attitudes toward children and school work prior to internship and that these attitudes were modified only slightly upon completion of the internship.

LaBue (1959) concluded that teacher attitudes as measured by the *Minnesota Teacher Attitude Inventory* (*MTAI*) were related significantly to knowledge of the principles of educational psychology, child development, and child behavior, but were not related significantly to objectivity, friendliness, emotional stability, personal relations, and other traits as measured by the *Guilford-Zimmerman Temperament Survey*.

Chandler (1959) and Mathis (1959) reported on the relation between types of salary schedules (merit *versus* nonmerit) and teacher morale; they concluded that, although morale appeared to be a general function of a multitude of interrelated variables, the type of pay plan in a school system was not a significant variable by itself in determining the morale level.

Two investigations were concerned with attitudes of teachers early in their careers. Making use of the *MTAI*, Rabinowitz and Rosenbaum (1960) reported a decline in scores of elementary school teachers before teaching and after three years experience. The scale was designed to measure attitudes conducive to those interpersonal relations that are satisfying to both teachers and pupils. The decline in scores, which was associated with geographical location, was greatest among teachers in New York City. However, the decline was not associated with the difficulty of the school as shown by mean IQ or economic status of pupils. Compared with initial scores, the data scored indicated less concern with pupil freedom and more concern with classroom order and academic standards.

Mason (1961) found a generally high level of satisfaction with teaching among beginning teachers, who were most contented with their social relations with fellow teachers, superiors, students, and parents, and least pleased with various aspects of salary, particularly when their salaries were compared with earnings of other occupations open to persons with the same level of education.

The National Education Association, Research Division (1963a) reported on the percepts of teachers who would certainly or probably rechoose or not choose teaching again. As was noted by Mason (1961) in his study of beginning teachers, this investigation revealed that more women and elementary school teachers than men and secondary school teachers expressed satisfaction with teaching.

In order to develop a profile that would depict a school's organizational climate, Halpin and Croft (1962) asked elementary school teachers and principals to what extent the behavior described in each item of an Organizational Climate Description Questionnaire characterized their schools. Halpin and Croft discriminated six organizational climates and put special emphasis on the concept of authenticity of behavior.

Reporting on faculty morale in Negro colleges, Thompson (1960) found that, while some factors reflected the unique position of the Negro colleges in American culture, a significant number of findings were in agreement with those of other investigations of teacher morale.

Summary of Research on Teacher Morale, Opinions, and Attitudes

A review of the research between 1958-63 pointed to the following major conclusions:

1. Morale is a general function of a multitude of interrelated variables and dimensions rather than a function of one or more isolated variables. However, the absence of adequate instruments that are anchored to a comprehensive theoretical conceptualization of morale limits the extent to which research can be done effectively.
2. The immediate supervisor or administrator is extremely important to a teacher's morale. Democratic administration can offset the effects of other factors that tend to produce low morale.
3. Congruity or lack of congruity of perceptions and expectations of school boards and teachers is extremely important to teacher morale.
4. Administrators and teachers frequently have a very different view of the level of morale and of what is important to teachers' morale. The greater the discrepancy between their expectations, the lower the morale.
5. Preparation programs for teachers are inadequate to the extent that they either develop or fail to change unrealistic attitudes about teaching—a circumstance that may result in disorientation for beginning teachers as well as in reduced morale.
6. Research needs to be done on the relation of morale to teacher performance and to personnel policies and practices.

Additional References: Bentley and Rempel (1963); Blocker and Richardson (1963); Thorndike and Hagen (1961).

Bibliography

ADAMS, JAMES F. *Problems in Counseling.* New York: Macmillan Co., 1962. Chapter 27, "Ethical Responsibilities of the Counselor in the Counseling Process," pp. 148-61.

ADVISORY COMMISSION ON INTERGOVERNMENTAL RELATIONS. *Transferability of Public Employee Retirement Credits Among Units of Government.* Report A-16. Washington, D.C.: the Commission, 1963. 92 pp.

ARMSTRONG, W. EARL, and STINNETT, T. M. *A Manual on Certification Requirements for School Personnel in the United States.* 1961 edition. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1962. 219 pp.

BAKKEN, CLARENCE J. *The Legal Basis for College Student Personnel Work.* Washington, D.C.: American College Personnel Association, a division of American Personnel and Guidance Association, 1961. 55 pp.

Belok, Michael, and Dowling, Fred. "The Teacher Image and the Teacher Shortage." *Phi Delta Kappan* 42: 255-56; March 1961.

Bolmeier, E. C. "Legal Scope of Teachers' Freedoms." *Educational Forum* 24: 199-206; January 1960.

Bryson, Joseph E. *Legality of Loyalty Oath and Non-Oath Requirements for Public School Teachers.* Doctor's thesis. Durham, N.C.: Duke University, 1961. 244 pp. Abstract: *Dissertation Abstracts* 22: 3898-99; No. 11, 1962.

Buck, Roy C. "The Extent of Social Participation Among Public School Teachers." *Journal of Educational Sociology* 33: 311-19; April 1960.

Carlson, Richard O. "Variation and Myth in the Social Status of Teachers." *Journal of Educational Sociology* 35: 104-18; November 1961.

Chandler, B. J. "Salary Policies and Teacher Morale." *Educational Administration and Supervision* 45: 107-10; March 1959.

Chilcott, John H. "The School Teacher Stereotype: A New Look." *Journal of Educational Sociology* 34: 389-90; May 1961.

Day, Harry P. "Attitude Changes of Beginning Teachers After Initial Teaching Experience." *Journal of Teacher Education* 10: 326-28; September 1959.

DeVita, James, and Kaczkowski, Henry. "Teaching as a Vocational Choice." *Educational Administration and Supervision* 45: 83-85; March 1959.

Drabick, Lawrence W. "The Teacher's Day: Analysis of Professional Rôle Perceptions." *Educational Administration and Supervision* 45: 329-36; November 1959.

Furness, Edna Lue. "The Image of the High School Teacher in American Literature." *Educational Forum* 24: 457-64; May 1960.

Garber, Lee. *The Yearbook of School Law 1962.* Danville, Ill.: Interstate Printers and Publishers, 1962. 234 pp.

Geis, Gilbert, and Huston, Robley. "Trends in the Dismissal of Tenure Teachers." *The Yearbook of School Law 1962.* Danville, Ill.: Interstate Printers and Publishers, 1962. pp. 199-209.

Groff, Patrick J. "The Social Status of Teachers." *Journal of Educational Sociology* 36: 20-25; September 1962.

Halpin, Andrew W., and Croft, Don B. *The Organizational Climate of Schools.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 543. Salt Lake City: University of Utah, 1962. 187 pp.

Hamilton, Robert R., and Mort, Paul R. *The Law and Public Education.* Second edition. New York: Foundation Press, 1959. Chapter 8, "Liability of School Districts, Officers and Employees," pp. 279-356; Chapter 9, "The Employment of Teachers," pp. 357-93; Chapter 10, "Discharge of Teachers," pp. 394-448; Chapter 11, "Teacher Tenure and Retirement Legislation," pp. 449-505.

Harap, Henry. "Morale." *Nation's Schools* 63: 55-57; June 1959.

Kershaw, Joseph A., and McKean, Roland N. *Teacher Shortages and Salary Schedules.* RAND Corporation Research Study. New York: McGraw-Hill Book Co., 1962. 203 pp.

Kigin, Denis J. *Teacher Liability in School-Shop Accidents.* Ann Arbor, Mich.: Prakken Publications, 1963. 128 pp.

Krislov, Joseph. *A Survey of State and Local Government Retirement Systems.* U.S. Department of Health, Education, and Welfare, Social Security Administration, Bureau of Old-Age and Survivors Insurance, Division of Program Analysis, Economic Studies Branch. Washington, D.C.: Government Printing Office, 1962. 47 pp.

Kuhlen, Raymond G. *Career Development in the Public School Teaching Profession with Special Reference to Changing Motivations, Pressures, Satisfactions and Dissatisfactions.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 048. Syracuse, N.Y.: Syracuse University, 1959. 213 pp.

Kuhlen, Raymond G., and Dipboye, Wilbert J. *Motivational and Personality Factors in the Selection of Elementary and Secondary School Teaching as a Career.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 047. Syracuse, N.Y.: Syracuse University, 1959. 130 pp.

LaBue, Anthony C. "Teachers' Classroom Attitudes." *Journal of Teacher Education* 10: 433-34; December 1959.

Ligtenberg, John. "Courts Hold Tenure Improves School System." *American Teacher Magazine* 43: 13-14, 20; October 1958.

Lindsey, Margaret, editor. *New Horizons for the Teaching Profession.* Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961. 243 pp.

Lipset, Seymour M., and Bendix, Reinhard. *Social Mobility in Industrial Society.* Berkeley: University of California, 1959. 309 pp.

Mason, Ward S. *The Beginning Teacher: Status and Career Orientations.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 644. Washington, D.C.: Government Printing Office, 1961. 196 pp.

Mathis, Claude. "The Relationship Between Salary Policies and Teacher Morale." *Journal of Educational Psychology* 50: 275-79; December 1959.

National Education Association, Division of Audiovisual Instructional Service. *Proceedings of the Conference on Professional Rights and Responsibilities of Teachers in Relation to the Newer Educational Media.* Washington, D.C.: the Association, 1962. 153 pp.

National Education Association, National Council on Teacher Retirement. *Proceedings of the Fortieth Annual Meeting.* Washington, D.C.: the Association, 1962. 116 pp.

National Education Association, Research Division. *Public-School Teachers and Collective Bargaining.* Washington, D.C.: the Association, 1958. 20 pp.

National Education Association, Research Division. *The Teacher and the Law.* Research Monograph 1959-M3. Washington, D.C.: the Association, 1959. 90 pp.

National Education Association, Research Division. *A Decade of Court Decisions on Teacher Retirement, 1950-1959.* Washington, D.C.: the Association, 1960. 50 pp.

National Education Association, Research Division. *Salaries and Salary Schedules of Urban School Employees, 1960-61.* Research Report 1961-R17. Washington, D.C.: the Association, 1961. 91 pp. (a)

National Education Association, Research Division. *Statistics of Retirement Systems to Which Teachers Belong.* Research Report 1961-R16. Washington, D.C.: the Association, 1961. 59 pp. (b)

National Education Association, Research Division. *Statutory Provisions and Judicial Decisions Relating to Political Activities of Public-School Teachers.* Research Memo 1961-32. Washington, D.C.: the Association, 1961. 11 pp. (c)

National Education Association, Research Division. *The Tax-Sheltered Annuity.* Retirement Income Series, No. 22. Washington, D.C.: the Association, 1961. 11 pp. (d)

National Education Association, Research Division. *Why Have Merit Plans for Teachers' Salaries Been Abandoned?* Research Report 1961-R3. Washington, D.C.: the Association, 1961. 51 pp. (e)

National Education Association, Research Division. *Classroom Teacher Salary Schedules, 1962-63, Districts Having 6,000 or More Pupils.* Research Report 1962-R11. Washington, D.C.: the Association, 1962. 112 pp. (a)

National Education Association, Research Division. *The Economic Status of Teachers in 1961-62.* Research Report 1962-R7. Washington, D.C.: the Association, 1962. 54 pp. (b)

National Education Association, Research Division. *Estimates of School Statistics, 1962-63.* Research Report 1962-R13. Washington, D.C.: the Association, 1962. 32 pp. (c)

National Education Association, Research Division. *High Spots in State School Legislation, January 1—August 1, 1962.* Research Report 1962-R9. Washington, D.C.: the Association, 1962. 38 pp. (d)

National Education Association, Research Division. *Professional Negotiation with School Boards.* Washington, D.C.: the Association, 1962. 32 pp. (e)

National Education Association, Research Division. *State Minimum-Salary Laws and Goal Schedules for Teachers, 1962-63.* Washington, D.C.: the Association, 1962. 45 pp. (f)

National Education Association, Research Division. *Tenure and Contracts.* School Law Summaries. Washington, D.C.: the Association, November 1962. (g)

National Education Association, Research Division. *The American Public-School Teacher, 1960-61.* Research Monograph 1963-M2. Washington, D.C.: the Association, 1963. 112 pp. (a)

National Education Association, Research Division. *Salary Schedule Maximums for Administrators, 1962-63, School Systems Having 6,000 or More Pupils.* Research Report 1963-R2. Washington, D.C.: the Association, 1963. 108 pp. (b)

NATIONAL EDUCATION ASSOCIATION, RESEARCH DIVISION. *The Teacher's Day in Court: Review of 1962*. Research Report 1963-R5. Washington, D.C.: the Association, 1963. 53 pp. (c)

NATIONAL EDUCATION ASSOCIATION, RESEARCH DIVISION. *Twenty-First Biennial Salary Survey of Public-School Employees, 1962-63*. Research Report 1963-R7. Washington, D.C.: the Association, 1963. Part A, 43 pp. Part B, 76 pp. Part C, 90 pp. (d)

NATIONAL EDUCATION ASSOCIATION, RESEARCH DIVISION. *Who Is Liable for Pupil Injuries?* Washington, D.C.: National Commission on Safety Education, National Education Association, 1963. 71 pp. (e)

O'DOWD, DONALD D., and BEARDSLEE, DAVID C. "The Student Image of the School Teacher." *Phi Delta Kappan* 42: 250-54; March 1961.

RABINOWITZ, WILLIAM, and ROSENBAUM, IRA. "Teaching Experience and Teachers' Attitudes." *Elementary School Journal* 60: 313-19; March 1960.

REDEFER, FREDERICK L. "Factors That Affect Teacher Morale." *Nation's Schools* 63: 59-62; February 1959.

REDEFER, FREDERICK L. *The Morale of Teachers*. Research Memo 1963-18. Washington, D.C.: Research Division, National Education Association, 1963. 21 pp.

REMMLEIN, MADALINE KINTER. *School Law*. Second edition. Danville, Ill.: Interstate Printers and Publishers, 1962. Part I, pp. 3-192.

REMMLEIN, MADALINE KINTER, and WARE, MARTHA L., editors. *An Evaluation of Existing Forms of School Laws*. Cincinnati: W. H. Anderson Co., 1959. Chapter 4, "Evaluation by Subject Matter," pp. 45-82.

ROETHEL, DAVID A. H. "1960 Chemical Salaries." *Chemical and Engineering News* 39: 118-24, 126; December 1961.

THOMPSON, DANIEL C. "Problems of Faculty Morale." *Journal of Negro Education* 29: 37-46; Winter 1960.

U.S. DEPARTMENT OF COMMERCE, OFFICE OF BUSINESS ECONOMICS. *U.S. Income and Output*. Supplement to *Survey of Current Business*. Washington, D.C.: Government Printing Office, July 1962. p. 29.

U.S. DEPARTMENT OF LABOR, BUREAU OF LABOR STATISTICS. *National Survey of Professional, Administrative, Technical, and Clerical Pay, Winter 1961-62*. Bulletin 1346. Washington, D.C.: Government Printing Office, 1962. 57 pp.

U.S. DEPARTMENT OF LABOR, WOMEN'S BUREAU. *Economic Indicators Relating to Equal Pay, 1962*. Pamphlet No. 9. Washington, D.C.: Government Printing Office, 1962. 19 pp.

U.S. DEPARTMENT OF LABOR, WOMEN'S BUREAU. *1962 Handbook on Women Workers*. Bulletin No. 285. Washington, D.C.: Government Printing Office, 1963. 202 pp.

WILLIAMSON, JAMES. "Court Calls Expert Panel To Testify on Teacher Conduct." *CTA Journal* 56: 15-16; March 1960.

Additional References

BENTLEY, RALPH R., and REMPEL, AVERNO M. "Peer-Selection vs. Expert Judgment as a Means of Validating a Teacher Morale Measuring Instrument." *Journal of Experimental Education* 31: 233-40; March 1963.

BLOCKER, CLYDE E., and RICHARDSON, RICHARD C. "Twenty-Five Years of Morale Research: A Critical Review." *Journal of Educational Sociology* 36: 200-10; January 1963.

CHARTERS, W. W., JR. "The Social Background of Teaching." *Handbook of Research on Teaching*. (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 14, pp. 715-813.

GAUERKE, WARREN E. *Legal and Ethical Responsibilities of School Personnel*. Englewood Cliffs, N.J.: Prentice-Hall, 1959. 302 pp.

THORNDIKE, ROBERT L., and HAGEN, ELIZABETH. "Men Teachers and Ex-Teachers: Some Attitudes and Traits." *Teachers College Record* 62: 306-16; January 1961.

WILSON, W. CODY, and GOETHALS, GEORGE W. "The Relation Between Teachers' Backgrounds and Their Educational Values." *Journal of Educational Psychology* 51: 292-98; October 1960.

CHAPTER VII

Assessment of Teacher Behavior and Instruction

DAVID G. RYANS

THERE HAS BEEN a large quantity of research reports dealing with teacher behavior during the five-year interval since the topic was last covered in the REVIEW. A wide range of variables has been investigated. An overview of the reports suggests that investigations have been piecemeal and fragmentary for the most part; only occasional attempts have been made to conduct coordinated programmatic research or to relate studies to some theoretical model.

Overview of Trends in Research on Teacher Behavior

In the reviewer's opinion, several general trends are in evidence:

1. Research on teacher behavior appears to have improved in quality in recent years. The state of the art seems to have changed since 1952, when, in introductory remarks to the June issue of the REVIEW, "Teachers and Nonacademic Personnel," the present reviewer deplored the scarcity of well-designed research and the relative dearth of significant investigations. Although 10 years later poorly designed research continues to be reported, a larger proportion of the output seems to meet reasonable standards. Moreover, many of the studies deal with problems of basic importance. Assuming the validity of the reviewer's opinion that a change for the better has taken place, one is tempted to surmise the influences involved. It would be difficult to assess the reasons with any assurance; nevertheless, one does have the impression that generally more sophisticated researchers are being granted advanced degrees, that there is improved understanding of research design and methodology, and that more efforts are being extended toward relating empirical results to organized theoretical positions. Undoubtedly, one influence for improved research has issued from federal support through the Cooperative Research Program and the National Defense Education Act Title VII Program of the U.S. Office of Education.

2. A second trend, which is undoubtedly related to the improvement in the quality of research, has been an increase in attention given to methodological problems. The criterion problem is no longer neglected as much as it was a few years ago; methods for observation and data collection have been refined and improved; and relatively sophisticated design and analysis techniques increasingly have been employed.

3. A third general trend has been a lessening of attention to the topic of teacher effectiveness, which often has been elusive in its ramifications because of the incomparability of evaluative criteria. Instead, one may note an increase in attention to the fundamental problem of the description of teacher behavior and to the conditions with which various teacher behaviors are correlated. One serious impediment to the understanding of teacher behavior has been the frequent failure of researchers to recognize that before progress can be made in any field of science there must be a long period of investigation devoted to (a) the identification of the elements or components that make up the phenomena under consideration and (b) the study of their interrelationships—the descriptive phase of research. For many years, there appeared to be little recognition that study and assessment of teacher effectiveness involved the adoption of a value system and that value judgments necessarily were relative in nature. The estimation of teacher effectiveness must be preceded by (a) description of teacher behaviors, (b) subjective judgment and choice of values and associated teacher behaviors that are acceptable to a particular group as manifestations of teacher effectiveness, and (c) well-designed study of the relationships between observable teacher characteristics and teacher behaviors on the one hand and the accepted evaluative criteria on the other. The current trend to seek improved understanding of teacher behavior rather than to move directly to the study of teacher effectiveness is commendable.

4. There also has been a significant trend—not yet of great magnitude but certainly gaining momentum—toward concern with the theory of instruction and with models of teacher behavior that may provide organizing principles for research. This theoretical orientation should exert an important influence on future research concerning teacher behavior.

In surveying the research, the reviewer has been forced by the volume of the literature to mention only a relatively small number of the studies that have been conducted. This vast quantity of research has necessitated the reviewer's assuming the role more of reporter than of critic. Only the studies that have appeared to be at least fairly adequate are reviewed; undoubtedly, some well-qualified ones have been omitted. It also is recognized that each research investigation conducted probably is subject to some valid criticism. The reviewer, nevertheless, chose not to resort to petty criticisms. In fact, a critique of individual studies will be largely omitted from the review.

General References

A number of volumes and special issues of journals dealing with teacher behavior have appeared recently or are scheduled for publication within the next few months. The reader should be aware of these.

The third edition of the *Encyclopedia of Educational Research* carried articles by Mitzel (1960) and Ryans (1960b) devoted to the topic of teacher effectiveness. Mitzel described important attributes of criteria of teacher effectiveness as they must be considered by the researcher and discussed the problem from three aspects: (a) product criteria, (b) process criteria, and (c) presage criteria. Ryans also emphasized how important the criterion problem is when one conducts research on the prediction of teacher effectiveness. He reviewed the current status of prediction research. His presentation was organized about topics that were concerned with the determination of estimates of the criterion, the choice of estimates of predictors, approaches to predictor-criterion relationships, tentative generalizations regarding the predictability of teacher effectiveness, and probable correlates of teacher effectiveness.

A second major contribution, which brought together massive amounts of information about education, was the *Handbook of Research on Teaching*, edited by Gage (1963a). This important volume consisted of four major sections of from three to nine chapters each. Part I was concerned with a theoretical orientation; Part II, with methodologies in research on teaching; Part III, with major variables and areas of research on teaching; and Part IV, with research on teaching various grade levels and subject matters. The volume provided an excellent bibliographical source. Of particular relevance to the study of teacher behavior were chapters by Charters (1963), Gage (1963b), Getzels and Jackson (1963), Medley and Mitzel (1963), Stern (1963), and Wallen and Travers (1963).

The American Council on Education published a relatively comprehensive volume, *Characteristics of Teachers* (Ryans, 1960a). In this book, Ryans brought together many of the results of a coordinated program of research which involved more than 6,000 teachers, about 1,700 schools, and approximately 450 school systems. The staff of the Teacher Characteristics Study conducted approximately 100 research projects—many consisting of multiple research studies.

Representative of the over-all effort were the following projects and topics discussed in the volume: (a) development and analysis of a Classroom Observation Record and accompanying Glossary; (b) study of the reliability of observations of the classroom behavior of teachers; (c) longitudinal investigations of observed teacher behavior; (d) factor analyses of observed classroom behavior of elementary and secondary teachers; (e) statistical analyses of educational points of view held by teachers; (f) development of and subsequent factor analyses of scales estimating teachers' attitudes toward groups contacted in the schools; (g) investigation of critical incidents in teaching; (h) development of a Teacher Characteristics Schedule; and (i) response analyses leading to Schedule scoring keys. These keys were used to obtain estimates of the following: (a) characteristics of the classroom behavior of teachers—e.g., warm, understanding, systematic, responsible, stimulating, imaginative; (b) teacher attitudes—such as those toward pupils or toward school personnel; (c) educational beliefs; (d) verbal understanding; and (e) emotional adjustment. For the Teacher Characteristics Schedule, comparisons were made of the scores of teachers who had been classified according to sex, age, marital status, length and kind of teaching experience, type of college attended, section of country in which they had been teaching, and size of community and of school in which their teaching had taken place.

Among several of the other topics considered were the following: (a) the relationships between teacher behavior and pupil behavior, (b) the criterion problem in studying teacher behavior, (c) the development of a theory in relation to teacher behavior, (d) the study of activities of teachers during childhood and adolescence, and (e) the description of characteristics of teachers who had been assessed uniformly high or uniformly low relative to classroom behaviors.

Many of the research studies described in the volume also appeared in considerable detail in journal articles which will not be cited here. The reader is referred to the bibliography of the over-all report (Ryans, 1963a).

The National Education Association, Research Division (1962b, c; 1963) produced three bulletins which provided data based on well-designed sampling surveys. These three bulletins, respectively, were concerned with furnishing information upon the following items: (a) personal characteristics—e.g., tabulations concerned with teacher's sex, marital status, education, experience, income and employment, duties, and certain attitudes toward teaching; (b) opinions about teaching-related topics—e.g., amount of academic work required of students, effect of television on pupils' study habits, advantages or disadvantages of ability grouping, advisability of revealing test scores to parents, and value of foreign language instruction in the elementary schools; and (c) miscellaneous facts relative to teacher's age, number and type of college degrees held, years of teaching experience, amount of income, teaching load in terms of size of classes, marital status and number of children in family, occupation of father, and participation in community activities.

Publications Containing Selected Research Papers

In a recent issue of the *Journal of Experimental Education,* edited by Barr (1961), approximately 75 doctoral studies were reviewed. Consideration was given to the criteria employed, to the importance of personal qualities of teachers, to the problems of motivation, and to the general trends and patterns of teacher effectiveness that were suggested by the investigations.

The April 1963 issue of *Theory into Practice*, edited by Jackson (1963), and the Spring 1963 issue of *School Review*, edited by Bidwell (1963), were devoted to a research-oriented presentation of problems related to teacher behavior and instruction. Another collection of papers with a strong research emphasis and with significant theoretical implications was an authoritative symposium on "Current Research on the Classroom Behavior of Teachers and Its Implications for Teacher Education," which at the time of the reviewer's writing was scheduled to appear in the September 1963 issue of the *Journal of Teacher Education.*

Theory of Instruction and Teacher Behavior

Theories and models are important not only for the organizing service they provide for empirical research but also for their equally important heuristic value and stimulation of further research. During recent years, several researchers have emphasized these points.

In urging that attention be given to theory, Gage (1963b) discussed the nature of theoretical models and described several different ones. Some of his models were taken from fields other than education—e.g., statistics, psychopathology, administrative behavior, group dynamics, and small-group research. Other models were directed specifically toward teaching-learning—e.g., paradigms concerned with criteria of effectiveness and with the teaching process. A strong case was made for a theory of teaching.

Perception-Diagnosis Theory of Teaching

B. O. Smith (1960) developed a concept of teaching which proposes that a teacher's perception of a pupil's behavior leads to a teacher's diagnosis of the pupil's state of interest, readiness, knowledge, and the like. The diagnosis is based on the teacher's inferences from the behavior of the pupil. In light of the diagnosis, the teacher takes action. At the same time, a pupil's perception of the teacher's behavior leads to a diagnosis on the part of the pupil of the teacher's state of interest—a diagnosis inferred from the teacher's behavior. Thus, the pupil reacts to the actions of the teacher. A cycling process arises that ultimately leads to pupil achievement if the goals of teaching are realized.

Teaching Conceptualized as a System of Social Action

Smith and Meux (1962) presented a concept of teaching as a system of social action involving an agent, an end-in-view, a situation, and two sets of factors in the situation. The agent is not in a position to control one set of factors—e.g., size of classroom, characteristics of pupils. However, the agent can modify the other set of factors with respect to the end-in-view—e.g., assignments, ways of asking questions. Modifiable conditions, which constitute the means by which ends-in-view are reached, consist of two types: (a) one concerning the subject matter and instructional materials and (b) the other concerning ways of manipulating the subject matter and materials. Moreover, the manipulation of the procedural features of teaching also has two major aspects: (a) large-scale maneuvers called *strategies*—referring to patterns of acts that serve to attain certain outcomes, such as insuring that certain learnings will be acquired in as brief a time as possible—and (b) smaller teacher-behavior components constituting tactical implements of strategies which are referred to as *logical operations*—forms which verbal behavior takes as the teacher shapes the subject matter in the course of instruction. Thus, logical operations are basic to the teaching process. A system was developed by Smith and Meux for classifying verbal teaching behavior into pedagogically significant units.

Information System Model

Ryans (1960a) extended a theoretical model first introduced in 1955-56 which emphasized the instrumental nature of teacher behavior, the effects of pupil interaction and feedback on teacher behavior, the aggregation of more elementary characteristics of teachers and the organization of these characteristics into molar teacher behaviors, the effect of situational conditions on teacher behavior at different levels of aggregation, and the systematic nature of teacher behavior. Subsequently, Ryans (1963a) elaborated upon this point of view. He suggested that instruction which properly may be viewed as a communication process may be described in terms of a theoretical model which incorporates the principles of an information system.

In the information system model, it is proposed that teacher information processing, which culminates in teacher behavior, is influenced by two major sets of inputs: (a) capabilities and characteristics of the teacher, or internal inputs—e.g., physical-physiological characteristics, general capabilities, characteristic abilities-capacities, characteristic behaving styles, characteristic affective sets, and retrievable information-behavior—and (b) conditions external to the teacher, or external inputs—e.g., objectives-goals of learning in a given situation, behavior content to be learned by the pupil, learning media available, pupil behavior, administrative policies, counseling to which the pupil has been exposed, and the culture of which the pupil is a part. Teacher information processing itself is held to consist of a series of logical phases tentatively classified as (a) input processing —sensing, filtering, analyzing and classifying, and temporary storing of inputs; (b) processing of information preliminary to decision making and channeling—goal defining and task analyzing, retrieving appropriate information-behavior, redintegrating information, and evaluating outcomes of possible alternate behaviors; (c) decision making—decisions concerning content of information to be conveyed, selection of mode of teacher behavior or channeling, and control of external conditions; (d) processing of information for use, or programming of teacher behavior; and (e) execution of the selected teacher behavior. Feedback occurs both from the teacher's behavior and from the pupil's behavior in a particular teaching situation. Feedback is hypothesized to influence teacher information processing and teacher behavior in future similar situations.

Ryans' information system theory maintains that in activities involved in the formal teaching-learning situation selected information in the form of facts, concepts, and rules relating to cognitive, affective, and psychomotor behavior is assembled, organized, and programed for conveyance via some medium—e.g., the teacher himself, a textbook, a TV display, or a programed lesson—to a receiver-learner. The manner in which this conveyance takes place is intended to enhance both accurate receipt of the information-behavior communicated and incorporation of the information into the learner's hierarchical behavior repertoire—an incorporation achieved by invoking active participation on the part of the receiver-learner. Communication in this context refers to the forwarding of meaningful information and, furthermore, to the conveyance of information that can be used by the receiver-learner. It involves the transmittal not only of signals comprising a message in the syntactic sense but also of references to the objects and events to which the employed symbols refer. Furthermore, it is dependent upon common past experiences of sender and receiver, on behavioral sets and states of readiness, and on current conditions and circumstances.

Model Viewing the Teacher as Problem Solver

Turner and Fattu (1960b) espoused a theoretical position in which the teacher is conceptualized primarily as a problem solver. The position taken is as follows: (a) teacher classroom behavior consists of responses that are instrumental in bringing about a goal in the form of pupil behavior held to be desirable by a particular teacher or group; (b) when the teacher sets his goal of bringing about a particular pupil behavior, he generates a problem for himself—that of getting the pupil to make the desired response and of making certain the pupil can produce that response when required; (c) effective teachers who are good solvers of teaching problems in a variety of teaching situations can transfer successfully their instrumental problem-solving responses to new or variant situations; (d) skill in solving teaching problems depends on ability to acquire instrumental responses and on the ability to utilize information in order to select the appropriate instrumental response in a given situation; (e) ability to acquire instrumental responses depends on the teacher's learning ability, on a climate of autonomy during early years of teaching, and on a breadth of goals learned; and (f) solving teaching problems can be estimated from constructed problems presented to the teacher—i.e., problems that appear to have high face validity, that differentiate teachers from nonteachers, that are solved more successfully by demonstration and supervisory teachers than by randomly selected teachers, and that ultimately can be shown to result in desired pupil behavior.

It may be noted that the points of view suggested by Ryans, Smith and Meux, and Turner and Fattu bear a resemblance to one another, at least to the extent that teacher behavior is assumed to consist of instrumental responses—those responses or behaviors having a logical, problem-solving, decision-making basis.

Center for the Construction of Theory in Education

A Center for the Construction of Theory in Education has been established at the Ohio State University. A number of papers dealing with the nature of theory and of models have been issued from this theory-oriented group. Descriptions of theoretical models for education have been developed by analogy and comparison with principles derived from different areas, such as agricultural economics, pharmacology and pharmacogenetics, and general systems theory. A recent paper by Maccia and Maccia (1962) was a representative one from the Center.

Methodology

Criterion Problem

Many research workers have discussed the criterion problem in studying teacher behavior and teacher effectiveness. Mitzel (1960) described important attributes of criteria of teacher effectiveness and used the terms *product, process*, and *presage* to classify criteria in terms of their proximity to educational goals. Barr (1961) reviewed criteria of teacher effectiveness employed in the Wisconsin Studies, including pupil gains, scores on instruments hypothesized to measure teacher effectiveness, self-ratings, ratings of teachers by other educational personnel, peer ratings, pupil ratings, and practice teaching marks. In addition to a discussion of the methodology relative to the development of predictors of teacher behavior and their validation, Ryans (1960a, b) considered such topics as the dimensionality of the criterion, sampling adequacy of the criterion, interrelationships of criterion dimensions, generalizability of criterion dimensions, requirements of adequate criterion measures—e.g., validity or relevance, reliability, feasibility—and approaches to criterion measurement. These approaches included not only direct measurement of the ongoing criterion behavior but also indirect measurement—i.e., an assessment either obtained from a preserved record of the behavior-in-process or derived from judgments of nontrained observers that were based on recall of past behavior or on evaluation of a product of the criterion behavior or on a quantifiable estimate of concomitants of criterion behavior.

In keeping with their point of view of teacher behavior as problem solving, Turner and Fattu (1960a) discussed various types of criteria related to effective problem solving: (a) ultimate criterion, which is complete and uncontaminated—for which data usually are unattainable; (b) consistency criteria, which are concerned with whether problem solving is compatible with information available for determining the solution of a problem and with whether the solution can be inferred from some type of empirical evidence, rationale, or theory; and (c) consensual criteria, which pertain to obtaining a consensus relative to what constitutes problem-solving behavior. A preference for the consistency criteria was indicated.

Classifying Teacher Behaviors

Medley and Mitzel (1962) suggested a three-dimensional scheme for classifying teacher behaviors as an aid to researchers. Teacher behaviors were classified as follows: A—the proportion of situations in which a behavior is effective; B—the proportion of pupils affected by the behavior; and C—the proportion of teachers for whom the particular behavior works effectively. The most effective teaching behaviors would be those high with respect to A, B, and C. Obviously, behaviors low on A, B, and C probably are not worth trying in teaching. The usefulness of studying different combinations of A-type, B-type, and C-type behaviors was discussed. In terms of defining the effectiveness of a particular behavior, it is important to consider the total number of teachers, pupils, and situations for which it is effective.

Observation and Assessment of Teaching

Medley and Mitzel (1958b, 1963) considered the observation and assessment of teaching. Medley and Mitzel (1958a, 1963) also discussed the reliability of the observation of teacher classroom behavior.

Remmers (1963) described various kinds of rating scales employed in teaching research and also treated sociometric methods, the semantic differential, Q techniques, and other rating procedures.

The National Education Association, Research Division (1962a) prepared a memorandum concerning Ryans' Classroom Observation Record and Glossary together with a summary of directions for their use. Ryans (1960a) described the Teacher Characteristics Schedule which was used in the Teacher Characteristics Study to estimate three major sets of teacher classroom behaviors as well as teacher attitudes, beliefs, verbal-semantic ability, and emotional adjustment. Although other instruments have been developed, those described are representative.

Flanders (1960a) provided a manual for the use of his technique of interaction analysis in the classroom (Flanders 1960b, 1961). The major innovation of this system was the use of a matrix which provides systematically arranged information about spontaneous classroom behaviors.

Bowers, Davis, and Bowers (1962) described their Index of Adjustment and Values which is intended to measure teacher attitudes toward self and others. Travers and others (1961) reported the use of an instrument for measuring teacher needs-motivation.

Interpersonal Relations in the Classroom

At least two studies have made use of pupil accomplices to study teacher behavior and the effects of teacher behavior upon other pupils. Keislar and McNeil (1959) employed pupils who were instructed to show pleasure under certain experimental conditions of individual spelling instruction (tutoring). In order to gain information about the effects of different teacher control techniques on classes and their reaction to the teacher, Kounin, Gump, and Ryan (1961) reviewed a number of studies in which students had played disruptive roles in the classroom and had evoked predetermined—that is, contrived—responses from the teacher. They recommended that researchers should be allowed to enter classrooms to study at first hand interpersonal relations of students and teachers.

Summary of Teacher-Student Variables Considered in the Literature

To provide a ready reference to teacher-student variables that have been considered in the literature of the past five years, a summary list has been prepared. For convenience, the list which constitutes an appendix to this chapter has been organized under these rubrics: (a) internal inputs to teacher information processing and behavior (physical-physiological characteristics, general capabilities of the human organism, characteristic abilities-capacities, characteristic teacher behaving styles, characteristic affective sets, and retrievable information); (b) external inputs to teacher behavior; (c) teacher information processing; (d) outputs of teacher information processing—teacher behaviors; and (e) products of teacher behaviors—pupil behaviors. (See Appendix on page 434.)

Internal Inputs Contributing to Instructor Behavior

Few studies have been made of the organic state, or the physical-physiological characteristics, and their relation to teacher behavior. Exceptions include studies about age by Ryans (1960a) and research carried out by Lucio and Wenger (1961) on some interesting hypotheses relating to autonomic functioning.

Studies of the capacities and abilities of teachers have been numerous over the years. (The category, general capabilities of the human organism, will not be discussed here, since these are, in a sense, capabilities for inputs rather than inputs *per se*. Moreover, there is no relevant literature available for review.) Only a few studies will be mentioned, although abilities estimated from test scores and college marks appeared frequently in reports that focused on other aspects of teacher behavior. Ryans (1960a) reported findings on both verbal-semantic ability and academic abilities of teachers. In a factor analysis of Q sorts of teacher classroom behavior, Wallen and others (1963) found "much academic activity versus little academic activity" as a major factor. Smith and Meux (1962) considered the abilities of teachers from the standpoint of logical operations that were observable in the verbal behavior of the teacher as the subject matter was presented.

Behaving styles of teachers have been studied extensively. In the reviewer's opinion, these should be considered primarily as underlying characteristics, or source traits, rather than teacher behaviors. Nevertheless, they also may be viewed as teacher behaviors when they occur in the classroom as the output or part of the output of teacher information processing and decision making in a particular teaching-learning situation. Although the behaving styles reported covered a wide range, certain styles appeared to have been observed with some regularity by different researchers. For example, a "friendly, warm, understanding, sensitive, sincere, empathic, rapport-maintaining" sort of style pattern was reported by Bowers, Davis, and Bowers (1962), Hodgson and Horst (1959), Medley (1962), Medley and Mitzel (1959), Nunnally, Thistlethwaite, and Wolfe (1963), Ryans (1960a), and Wallen and others (1963) as well as by other investigators during recent years. Similarly, a "responsible, systematic, businesslike, dependable, well-prepared, efficient-presentation, student-achievement-oriented, orderly, organized" pattern of style was reported by Heil and Washburne (1962), Hodgson and Horst (1959), Medley (1962), Nunnally, Thistlethwaite, and Wolfe (1963), Ryans (1960a), and Stern and others (1960). A pattern of styles involving "stimulating, imaginative, achievement-motivating, active, resourceful, dynamic" teacher characteristics was reported by Cotsonas and Kaiser (1962), Hodgson and Horst (1959), Medley (1962), Ryans (1960a), and Wallen and others (1963). A "personally acceptable, attractive, good appearing, verbally expressive, communicative, professionally impressing" style pattern was reported by Bowers, Davis, and Bowers (1962), Hodgson and Horst (1959), Medley and Mitzel (1958b), Nunnally, Thistlethwaite, and Wolfe (1963), Ryans (1960a), and Stern and others (1960).

Another type of pattern studied by a number of investigators had to do with a "direct, dominative, teacher-centered, authoritarian" style versus an "indirect, integrated, permissive, child-centered" style. Flanders (1960a, b; 1961), Stern and others (1960), and Wilk and Edson (1962) studied these patterns. As is apparent in the list presented in the Appendix, many other styles also have been noted in the research reports.

Affective sets of teachers have been studied extensively. One of the most frequently occurring patterns reported in the literature was concerned with what might be called traditional *versus* progressive orientation or teacher- and academic-centered *versus* child-centered orientation. Christensen (1960), Cook, Linden, and McKay (1961), Kerlinger and Kaya (1959), Nunnally, Thistlethwaite, and Wolfe (1963), Ryans (1960a), and Schmid and others (1962) investigated this dimension. Attitudes toward pupils were studied by Hoyt and Cook (1960), Ryans (1960a), and many other investigators who employed the *Minnesota Teacher Attitude Inventory*. Personal adjustment was explored by Ryans (1960a, c), Soar (1962), and Stern and others (1960); introversion-extroversion, by Cook, Linden, and McKay (1961). Role expectations of teachers were considered by Biddle, Twyman, and Rankin (1962), Rosencranz and Biddle (1962), L. M. Smith (1960), and Stern and others (1960).

The teacher's retrievable information, or knowledge of subject matter content, general culture, and teaching techniques, was involved in reports by Cosgrove (1959), Cotsonas and Kaiser (1962), Ryans (1960a), and Smith and Meux (1962).

External Inputs—Conditions External to the Teacher

With respect to external inputs to teacher information processing—inputs that come from the environment in which the teacher operates—the published studies have been somewhat limited, at least insofar as the reported research has indicated an attempt to relate environmental conditions of the teacher to teacher behavior. Wittrock (1962) studied the effectiveness of directions provided for the teacher. Ryans (1960a) attempted to establish relationships between (a) certain behaving styles and teacher characteristics on the one hand and (b) environmental conditions on the other—e.g., grade level and subject taught, size of school in which teaching occurred, size of community, socioeconomic status of community in which teaching took place, cultural level of community, and methodological emphasis of the school in which teaching was done.

Nunnally, Thistlethwaite, and Wolfe (1963) identified factors involved in the instructional environment. These factors were divided into two sets. One set was concerned with the faculty; and the other, with student conditions. The latter set included intellectual drive, competition, scientific interest, and pressure against scholarly activities.

Teacher Information Processing Behavior

Relatively few attempts have been made to infer what takes place within the "black box" when the internal inputs—i.e., characteristics of the teacher—and the external inputs—i.e., interacting conditions and subsystem characteristics external to the teacher—are manipulated within the teacher's cognitive system, both as the teacher is making decisions about how to behave in a specified teaching-learning situation and as the teacher behavior is being programed to convey to the pupil information appropriate to the teaching objective and behavior content to be learned. This inferential area is one in which both attempts to conceptualize and subsequent research to test the construct validity of propositions growing out of inferences about how the teacher processes information are needed.

As previously noted, Smith and Meux (1962) attempted to describe certain logical operations involved in teaching as the teacher generates or programs strategies—i.e., patterns of teacher behavior that serve to attain outcomes, such as insuring certain pupil learnings. As discussed in the previous section on theory of instruction and teacher behavior, Ryans (1963a, b) attempted to describe teacher information processing—i.e., publicly nonobservable processes involved as the teacher makes decisions and channels behavior to attain certain objectives—in terms of four hypothetical phases: (a) input processing, (b) processing of information preliminary to decision making and channeling, (c) decision making, and (d) processing of information output for use—i.e., programming of behavior—preparatory to conveyance of information to the pupil system. These phases are presumed to culminate in teacher behavior in a particular situation, or in what Smith and Meux referred to as teacher strategy employed in a particular teaching situation. Important to Ryans' formulation is the concept of negative feedback, which emphasizes the effect of a teacher's behavior and of a pupil's response to that behavior on future teacher behaviors in similar situations. The feedback information serves as an additional input for future teacher information processing.

Turner and Fattu (1960a) viewed teacher information processing essentially as problem solving—i.e., the solving of problems related to teacher situations. The views of Ryans, Smith and Meux, and Turner and Fattu are in no sense antithetical; instead, they overlap considerably.

One of the few empirical studies concerning the influences of feedback on teacher behavior was that of Gage, Runkel, and Chatterjee (1960), in which the effects of feedback from pupils to teachers was studied. Comparison of experimental and control groups of sixth grade teachers indicated that, when teachers were provided with information obtained from their pupils regarding how the pupils described their actual teacher and how they described their ideal teacher on 12 items of teacher behavior, (a) teachers' behaviors changed—as indicated by subsequent pupil descriptions of their actual teacher—in the direction of the pupils' initial description of their ideal teacher, and (b) teachers became increasingly accurate in predicting their pupils' descriptions of their teacher.

Teacher Behaviors—Outputs of Teacher Information Processing

Two major approaches to research on the outputs of teachers—i.e., teacher behaviors—have been through (a) the description of teacher behavior based on direct observation of acts of teachers in teaching situations and (b) the description of teacher behavior based on inferences arrived at through observation of the acts and achievements of pupils in the teachers' classes. The first approach describes teacher behavior in particular teaching situations in terms of a logical classification of acts that can be directly observed; the second assumes that certain pupil behaviors and estimated achievements reflect teacher behavior that may have occurred in a given situation.

Description of Teacher Behavior Based on Direct Observation

With regard to the description of directly observable outputs of teachers, various classifications have been suggested. In Ryans' (1963a, b) most recent conceptualization of the outputs of teacher information processing, five major categories of teacher behaviors, or strategies, were hypothesized: (a) presenting teacher behaviors; (b) motivating teacher behaviors; (c) organizing-coordinating teacher behaviors; (d) evaluating teacher behaviors; and (e) counseling-guiding teacher behaviors. No evidence is available for the validity of this conceptualization of teacher outputs. Furthermore, it is probable that in a given teaching situation teacher behavior may consist of some combination of the separate behaviors in these five categories.

Smith and Meux (1962) conceptualized the major outputs of the teacher and of the teacher's logical operations to be large-scale maneuvers which, as has been noted, are termed *strategies*. These outputs look toward the efficient acquisition and reinforcement of certain learnings in pupils.

Similarly, Turner and Fattu (1960a) considered teacher behavior to consist of acts that are instrumental in bringing about a goal state. This state is definable in terms of desirable pupil behavior. Teacher behavior is directed toward effecting a desired response by the pupil and toward making certain that the pupil can produce that response when required. Turner and Fattu attempted to infer the level of effectiveness of a teacher behavior from the extent to which problem-solving tests differentiated among teachers of different abilities and degrees of training and experience.

Heil and Washburne (1962) classified teachers as type A (turbulent and impulsive), type B (self-controlling, orderly, and work-oriented), and type C (fearful). Type B teachers were found to be higher than type A or type C with respect to demonstrating leadership, identifying with others and helping others, exhibiting dramatic behavior, and showing methodical behavior. However, type B teachers were lower than the other groups with respect to humor, acceptance of impulses, and aggression.

Description of Teacher Behavior Based on Inferences

Various investigators have inferred teacher behavior from pupil behavior and/or pupil achievement. Medley and Mitzel (1959) supplied evidence to support the proposition that pupil-teacher rapport indicated the warmth, friendliness, and understanding behavior of the teacher. Kounin and Gump (1961) reported pupil aggression and misconduct to reflect punitive teacher behavior.

Two investigations involving increments in pupil learning pointed to a possible inference about the nature of teacher behavior. Heil and Washburne (1962) reported that pupils' gains on the *Stanford Achievement Test*, pupils' positive feelings, and children's social acceptance of one another suggested self-controlling, orderly, and work-oriented behavior of the teacher. Flanders (1960b) found increased pupil learning to be greatest in the classes of teachers who displayed flexible behavior—i.e., behavior that reflected predominantly direct or dominative teacher influence in certain situations, as when the student's perception of the goal was clear and acceptable, and predominantly indirect teacher influence in other situations.

Ryans (1961a, b) reported differences between the elementary and secondary school in the extent to which so-called productive pupil behaviors reflected subpatterns of teacher behavior. Relationships between observed teacher behaviors and observed pupil behaviors were relatively high in the elementary school but relatively low in secondary school classes.

In one of his recent papers, Medley (1962) described teacher behavior both in terms of directly observable acts of the teacher and in terms of reflections of pupil behavior. Eight major dimensions of teacher classroom behavior were suggested: (a) activity of pupils—teacher behavior eliciting a large amount of pupil response; (b) presence—teacher behavior producing high classroom order and favorable impression upon observer; (c) initiative—teacher behavior permitting a high degree of pupil initiative in determining class activities; (d) informative teaching—teacher behavior emphasizing objectives, information, clarification, and summarization; (e) imaginative teaching—teacher behavior stressing individualization and use of interesting and original devices and materials; (f) consideration—teacher behavior showing an awareness and consideration of pupils' problems and feelings; (g) warmth—teacher behavior exhibiting supportive qualities; and (h) disapproval—teacher behavior including frequent reproof and criticism of pupils.

Products of Instructor Behaviors: Pupil Behaviors

A substantial number of studies dealt directly with pupil behaviors. In most instances, the assumption has been implicit that the pupil behaviors were an effect or consequence of teacher behaviors. Most of the studies tested hypotheses regarding the influence of variously estimated teacher variables and variously defined and measured pupil variables.

Bryan (1963) reported a high correlation between the development of interest and liking for subject matter on the part of chemistry students on the one hand and teacher prestige on the other, as judged by students' ratings. Bryan also found that high school teachers can effectively use information gained from student-reaction feedback as a means of improving their image with students. Working with student teachers, Wilk and Edson (1962) discovered that, when the learning climate was characterized by supportive-encouraging teacher behavior, when a variety of teaching materials and procedures were employed, and when frequent pupil activity was evident, there were fewer disruptive pupil behaviors.

Pupil Achievement and Learning

Christensen (1960) studied pupil achievement and pupil affect-need in relation to subpatterns of teacher behavior. He reported that teacher permissiveness was unrelated to pupil achievement but that teacher warmth was significantly related to pupil growth in vocabulary and arithmetic. As already noted, Flanders (1960b) found that indirect teacher behavior fostered increased pupil learnings but that the pupils who learned most and scored highest on classroom attitude scales were those exposed to flexible patterns of teacher influence. Heil and Washburne (1962) noted that pupil gains on the *Stanford Achievement Test* were greatest with self-controlling teachers, intermediate with turbulent teachers, and least with fearful teachers.

Medley and Mitzel (1959) reported nonsignificant differences for reading growth and problem-solving skill of pupils studied in relation to teacher warmth-friendliness, teacher verbal emphasis, and teacher acceptance of pupil autonomy. Compared with the modest degree of association obtained between verbal emphasis of the teacher and pupil-teacher rapport, the extent of the relationship between pupil-teacher rapport and teacher warmth-friendliness was defined more clearly.

Turner and others (1963) reported on a series of studies associated with the thesis that effective teachers are persons who are good solvers of teaching problems in a wide variety of teaching situations. The results from one of the studies indicated that pupils taught by high-scoring teachers—good problem solvers—in grades four and five attained significantly greater achievement than pupils taught by low-scoring teachers in the same grades.

Page (1958) found a measurable effect on student effort and improvement in learning when secondary school teachers wrote comments, believed by them to be "encouraging," on student papers when the papers were returned. Rosenfeld and Zander (1961) determined that the extent to which students set their aspired grades was congruent with their perceived capacities. The observed relationship appeared to be influenced by (a) teacher disapproval when a student was performing as well as he felt that he could, (b) indiscriminate reward, and (c) legitimate reward for adequate performance, with accompanying increase in aspiration level.

In working with student teachers, Wittrock (1962) found that those who were told that their final course grades would depend on the amount of pupil gain in achievement produced actually did produce significantly more pupil gain than did a control group of student teachers who did not receive the set involving reward for pupil gain.

Teacher Effectiveness

Ryans (1960b) summarily reviewed the problem of the prediction of teacher effectiveness. However, as noted in the introductory remarks to this chapter, the number of studies concerning teacher effectiveness has diminished during the past few years. The philosophy involved will not be reiterated here.

In an intensive study of groups of teachers who had been uniformly assessed as high or low with respect to warm-friendly teacher behavior, organized-businesslike teacher behavior, and stimulating-imaginative teacher behavior, Ryans (1960a) noted a number of characteristics which were uniquely associated with the high as compared with the low and with the low as compared with the high groups.

Torrance (1960) analyzed the daily logs of activities of teachers. Those teachers who were judged from their logs to be most effective and least effective were compared with respect to cognition, memory, convergent thinking, divergent thinking, and evaluation. The more effective teachers did tend to report more thinking activities—convergent, divergent, and evaluative—than did the less effective ones. On the other hand, the less effective teachers did cite more recognitive and memory experiences than did the more effective ones.

Approaches to the merit rating of teachers have attracted the attention of many school people. Burch and Ferguson (1961) described a sophisticated approach employed in connection with the administration of the West Hartford merit rating of teachers.

College Instruction

McKeachie has been a leader in conducting and reviewing research on teaching and teacher behavior at the college and university level. In his chapter in the *Handbook of Research on Teaching*, McKeachie (1963) considered in some detail research in the area of college and university teaching. A bibliography of more than 200 titles was provided.

Limitations of Research Findings

In a review of this nature, it is appropriate to note certain limitations involved in the research findings reported as well as restrictions on the application of the findings. Several will be mentioned.

The results of studies reviewed have pertained to identifiable patterns of teacher behavior and to related topics. These results reflect contemporary behaviors of teachers. Regardless of how valid the results may be for the time at which the studies were conducted, the findings for replicated investigations may be quite different at another time or, for that matter, in another cultural setting. One may generalize in time or from culture to culture only when conditions are approximately constant. In society, change often is rapid. Conditions sometimes vary widely over geographical space. It is important to note that what has been recorded may or may not be generalizable to the future or to different cultures.

The data obtained and the findings reported relative to teacher characteristics and teacher behavior do not indicate what value decisions should be formulated. The results of these research studies not only are descriptive of particular samples of teachers but also are variable relative to the cultural milieu which provides the setting at the time and place of the investigation.

It also must be recalled that one hazard involved in the utilization of most behavioral research findings rests in their application to individual cases. For a large group of teachers, considered as a group and viewed actuarially, a principle of teacher behavior may be stated with rather high probability. For the same behavior, one would be able to make such a statement or prediction with regard to one individual teacher with a probability perhaps only slightly higher than chance expectation and therefore with substantially reduced assurance.

Several studies have indicated that assembled descriptions of teacher behavior are not equally generalizable to all kinds of teachers. Just as the reliability and validity of tests or measurements of a certain psychological construct have come to be viewed as relatively specific in nature—i.e., limited to the particular instrument used, in a defined class of situations, for a defined purpose—so also must the results of teacher characteristics research be considered. Although some findings seem to possess considerable generality, it would be well to keep in mind that present behaviors probably vary in only partially known ways from one grade level to another, from one subject matter to another, and from one social context to another.

Most of the reported investigations of teacher behavior and teacher characteristics were of a taxonomic or descriptive sort. Although they are appropriate, descriptive observations *per se* do not provide explanations of teacher behavior. When interrelationships among teacher characteristics or among teacher characteristics and pupil characteristics are reported, they often are of a correlational nature. Moreover, antecedent-consequent relationships usually cannot be inferred. Longitudinal studies should be conducted to help clarify conceptualizations about such matters. However, with respect to most teacher characteristics and behaviors and with respect to the conditions that influence them, the current state of information does not permit the luxury of inferences about producer-product relationships.

Future Research on Teacher Behavior

In a recent article, Ryans (1963c) suggested a limited number of possible directions of future research. A consideration of the contents of the article and additional reflections by the reviewer pointed to the following courses of action that might be taken to give direction to future research: (a) taking account of shortcomings of past research; (b) analyzing areas of teacher behavior and the teaching-learning process; (c) evaluating recent trends in teacher behavior research; (d) surveying the changing social, economic, and educational conditions; (e) reviewing related methodological problems; and (f) considering the theoretical framework of teacher learning. Some of the directions of future research might include the following: (a) extension of the identification of properties of teacher behavior; (b) study of the relevant aggregates or compounds of teacher behavior that have importance for the teaching-learning process; (c) attention directed toward prediction which conforms closely with the "prediction research paradigm"; (d) methodological investigation oriented toward the development of tools and techniques for obtaining increasingly valid information about teacher behavior; (e) systematic inquiry aimed at the operational description of pupil behavior goals; (f) formulation of relationships between molar teacher behaviors and external or environmental inputs, as well as between molar teacher behaviors and the internal inputs or teacher characteristics; (g) determination of the effects of feedback to the teacher based on his previous behavior and of the results of feedback to the teacher on pupils; (h) exploration of how teacher behaviors vary and of how they relate to observable, operationally defined behaviors of atypical pupils, such as the gifted, the retarded, and the handicapped; (i) analysis of the influence of different kinds of teacher education programs and courses, of practice teaching and its variations, and of various other educational experiences on the developing patterns of teacher characteristics; (j) ascertainment of the influence of in-service experiences of teachers, both organized and informal, on patterns of teacher characteristics and behaviors; (k) completion of comparative studies of the characteristics of teachers of different national, political, and cultural backgrounds; (l) experimentation with the use of simulation in developing intended teacher behaviors during teacher education; (m) study of the programming of teacher behavior; (n) description and analysis of behaviors of the teacher operating as an organizer-coordinator of resources and technological media; (o) specification of the influence of automation and of flexible scheduling of students on teacher behavior; (p) analysis of the influence of social and economic changes on teacher behavior; (q) determination of how basic or elementary teacher behaviors aggregate to form patterns of teacher behavior; (r) development of a conceptual framework for understanding the interactions of various operational conditions with the various teacher behaviors necessary to produce certain pupil behaviors; (s) systematic theoretical and experimental inquiry regarding the nature of teacher input processing, decision making, output processing, and information control; and (t) investigation of the problems of implementing innovations and new developments into the teacher's repertoire of behavior.

Appendix

Summary of teacher-student variables referred to in representative reports and papers, 1958-63

Inputs to Teacher Information Processing

Internal Inputs—Teacher Characteristics

1. Physical-physiological characteristics (Lucio and Wenger, 1961; Ryans 1960a, 1963a)
2. General capabilities of human organism (Ryans, 1963a)
3. Characteristic abilities-capacities
    a. Academic ability (Wallen and others, 1963)
    b. Verbal-semantic ability (Ryans, 1960a)
4. Characteristic teacher behaving styles
    a. Achievement-motivating (Hodgson and Horst, 1959)
        Enthusiastic (Cosgrove, 1959)
        Imaginative (Medley, 1962)
        Inspiring, stimulating (Cotsonas and Kaiser, 1962)
        Stimulating (Hodgson and Horst, 1959)
        Stimulating, imaginative (Ryans, 1960a)
        Vigorous, dynamic (Wallen and others, 1963)
    b. Avoidance (Cook, Linden, and McKay, 1961)
        Preclusive behavior (Cogan, 1958a, b)
    c. Businesslike, responsible, systematic (Ryans, 1960a; Nunnally, Thistlethwaite, and Wolfe, 1963)
        Orderly (Stern and others, 1960)
        Organized (Hodgson and Horst, 1959)
        Planning and procedures emphasis (Cosgrove, 1959)
        Preclusive behavior (Cogan, 1958a, b)
        Presence (Medley, 1962)
        Self-controlling, orderly, work-oriented (Heil and Washburne, 1962)
    d. Conjunctive behavior (Cogan, 1958a, b)
    e. Consideration (Medley, 1962; L. M. Smith, 1960)
        Emotional climate—warmth, friendliness (Medley and Mitzel, 1959)
        Friendly, understanding (Ryans, 1960a; Nunnally, Thistlethwaite, and Wolfe, 1963)
        Pupil rapport-maintaining (Cosgrove, 1959)
        Sensitive, sincere (Hodgson and Horst, 1959)
        Sensitive, sympathetic (Cotsonas and Kaiser, 1962)
        Supportive (Bowers, Davis, and Bowers, 1962)
        Warmth (Christensen, 1960; Medley, 1962; Wallen and others, 1963)
    f. Critical (Stern and others, 1960)

- g. Direct—teacher-centered (Flanders 1960a, b; 1961)
  Dominant (Stern and others, 1960)
- h. Disapproval (Medley, 1962)
- i. Exhibitionistic (Stern and others, 1960)
  Expressive (Ryans, 1960a)
  Personally and academically impressive (Hodgson and Horst, 1959; Nunnally, Thistlethwaite, and Wolfe, 1963)
  Verbal emphasis (Medley and Mitzel, 1958b)
  Verbal output (Bowers, Davis, and Bowers, 1962)
- j. Fearful (Heil and Washburne, 1962)
- k. Flexible (Flanders, 1960a, b; 1961)
- l. Indirect—nondirective (Flanders, 1960a, b; 1961)
  Integrative (Wilk and Edson, 1962)
  Nondirective (Stern and others, 1960)
- m. Inclusive behavior (Cogan, 1958a, b)
- n. Informative (Medley, 1962)
  Initiating structure (L. M. Smith, 1960)
  Problem structuring statements (Bowers, Davis, and Bowers, 1962)
- o. Turbulent, impulsive (Heil and Washburne, 1962)

5. Characteristic affective sets
   - a. Academic-centered (Ryans, 1960a; Nunnally, Thistlethwaite, and Wolfe, 1963)
     Traditional orientation (Kerlinger and Kaya, 1959)
   - b. Achievement-motivated (Schmid and others, 1962)
     Status-striving (Stern and others, 1960)
   - c. Affiliation-motivated (Schmid and others, 1962)
   - d. Authoritarianism (Cook, Linden, and McKay, 1961)
   - e. Confident (Wallen and others, 1963)
   - f. Compulsive conformity (Cook, Linden, and McKay, 1961)
   - g. Dependency (Cook, Linden, and McKay, 1961)
     Dependent (Stern and others, 1960)
     Social dependency (Soar, 1962)
   - h. Docility (Cook, Linden, and McKay, 1961)
   - i. Emotional adjustment (Ryans, 1960a, c)
     Personal maladjustment (Soar, 1962)
     Preadult fixated (Stern and others, 1960)
   - j. Introversion-extroversion (Cook, Linden, and McKay, 1961)
   - k. Knowledge of pupil orientation (Schmid and others, 1962)
   - l. Nurturant (Stern and others, 1960)
   - m. Permissiveness (Christensen, 1960)
     Permissive—child-centered (Ryans, 1960a)
     Progressive orientation (Kerlinger and Kaya, 1959)
   - n. Professional attitude (Schmid and others, 1962)
     Professionalism (Soar, 1962)
     Professional satisfaction (Hodgson and Horst, 1959)
   - o. Punitive behavior (Kounin, Gump, and Ryan, 1961)
   - p. Professional-development-oriented (Schmid and others, 1962)
   - q. Pupil rapport orientation (Hoyt and Cook, 1960)
   - r. Pupil attitude—favorable (Ryans, 1960a)
   - s. Social superficiality (Soar, 1962)
   - t. Time spent alone (Wallen and others, 1963)
6. Retrievable information
   - a. Knowledge of content and techniques (Cotsonas and Kaiser, 1962)
     Subject matter knowledge and organization (Cosgrove, 1959)
     Subject matter knowledge (Ryans, 1960a; Smith and Meux. 1962)
   - b. Knowledge of professional information (Ryans, 1960a)

External Inputs

1. Directions to teacher (Wittrock, 1962)
2. Grade level or subject taught, size of school, size of community, socioeconomic status of community, cultural level of community, and methodological emphasis of school (Ryans, 1960a)
3. Instructional environment (Nunnally, Thistlethwaite, and Wolfe, 1963)
4. Student conditions: intellectual drive, competition, scientific interest, pressure against scholarly activities (Nunnally, Thistlethwaite, and Wolfe, 1963)

Teacher Information Processing

1. Information processing by teacher: filtering, redintegration, decision making, encoding, channeling, control of information—i.e., cognitive, affective, and psychomotor information (Ryans, 1963a, b)
2. Teacher feedback effects
   a. Effect of interaction analysis training (Flanders, 1961)
   b. Effects of pupils' description of own and ideal teacher (Gage, Runkel, and Chatterjee, 1960)
   c. Effects of student reactions (Bryan, 1963)
   d. Theory of feedback (Ryans, 1963a)
3. Teacher logical operations
   a. Classes of logical operations (Smith and Meux, 1962)
   b. Conditions influencing, classes of, and information flow of logical processing—theoretical (Ryans, 1963a)
4. Teacher problem solving (Turner and Fattu, 1960b)

Outputs of Teacher Information Processing—Teacher Behaviors

1. Strategies followed by teacher to accomplish pupil learning (Smith and Meux, 1962)
2. Teacher counseling behavior (Ryans, 1963a)
3. Teacher evaluating behavior
   a. Fair grading (Hodgson and Horst, 1959)
   b. Teacher evaluating behavior (Ryans, 1963a)
4. Teacher motivating behavior (Ryans, 1963a)
5. Teacher organizing-coordinating behavior (Ryans, 1963a)
6. Teacher presenting behavior (Ryans, 1963a)
   Teacher problem solving in teaching situations (Turner and Fattu, 1961)

Products of Teacher Behaviors—Pupil Behaviors

1. Effects on pupil effort and learning of feedback from teacher to pupil (Page, 1958)
2. Pupil achievement
   a. Pupil achievement (Turner and others, 1963)
   b. Pupil learning (Flanders, 1960b)
   c. Pupil gains (Christensen, 1960)

  d. Pupil reading growth (Medley and Mitzel, 1959)
  e. Pupil test gains (Heil and Washburne, 1962)
3. Pupil activity (Medley, 1962; Wilk and Edson, 1962)
4. Pupil alertness (Ryans, 1960a; 1961a, b)
5. Pupil aspirations (Rosenfeld and Zander, 1961)
6. Pupil conduct
   a. Pupil conduct in class (Kounin, Gump, and Ryan, 1961)
   b. Pupil disruptive acts (Wilk and Edson, 1962)
   c. Pupil opposition (Heil and Washburne, 1962)
7. Pupil conformance (Heil and Washburne, 1962)
8. Pupil confidence (Ryans, 1960a; 1961a, b)
9. Pupil expectations and motives (L. M. Smith, 1960)
10. Pupil feelings (Heil and Washburne, 1962)
11. Pupil initiating behavior
    a. Initiating behavior (Medley, 1962; Ryans, 1960a, 1961a, b)
    b. Self-initiated activities (Cogan, 1958a, b)
12. Pupil interest in subject matter (Bryan, 1963)
13. Pupil leadership activities (Bowers, Davis, and Bowers, 1962)
14. Pupil rapport (Medley and Mitzel, 1959)
15. Pupil problem-solving skill (Medley and Mitzel, 1959)
16. Pupil required activities (Cogan, 1958a, b)
17. Pupil responsibility (Ryans, 1960a, 1961a, b)
18. Pupil social acceptance of one another (Heil and Washburne, 1962)
19. Pupil striving (Heil and Washburne, 1962)
20. Pupil waivering (Heil and Washburne, 1962)
21. Social organization of class (Medley and Mitzel, 1958b)
22. Traditional activities of pupils (Bowers, Davis, and Bowers, 1962)
23. Verbal activities of pupils (Bowers, Davis, and Bowers, 1962)

Bibliography

BARR, A. S., issue editor. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations." *Journal of Experimental Education* 30: 5-156; September 1961.

BIDDLE, BRUCE J.; TWYMAN, J. PASCHAL; and RANKIN, EARL F. "The Role of the Teacher and Occupational Choice." *School Review* 70: 191-206; Summer 1962.

BIDWELL, CHARLES E., editor. "Teaching: A Profession and a Process." *School Review* 71: 1-111; Spring 1963.

BOWERS, NORMAN D.; DAVIS, O. L., JR.; and BOWERS, MARY W. "The Effectiveness of the Index of Adjustment and Values in Predicting Classroom Behavior." *Nineteenth Yearbook of National Council on Measurement in Education.* Ames, Iowa: National Council on Measurement in Education, 1962. pp. 112-20.

BRYAN, ROY C. *Reactions to Teachers by Students, Parents and Administrators.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 668. Kalamazoo: Western Michigan University, 1963. 58 pp.

BURCH, PAUL R., and FERGUSON, LEONARD. *Summary of Steps Taken To Adjust and Translate the Teacher Appraisal Information to Final Scores.* West Hartford, Conn.: West Hartford Public Schools, 1961. 12 pp. (Mimeo.)

CHARTERS, W. W., JR. "The Social Background of Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 14, pp. 715-813.

CHRISTENSEN, C. M. "Relationships Between Pupil Achievement, Pupil Affect-Need, Teacher Warmth, and Teacher Permissiveness." *Journal of Educational Psychology* 51: 169-74; June 1960.

COGAN, MORRIS L. "The Behavior of Teachers and the Productive Behavior of Their Pupils: I. Perception Analysis." *Journal of Experimental Education* 27: 89-105; December 1958. (a)

COGAN, MORRIS L. "The Behavior of Teachers and the Productive Behavior of Their Pupils: II. 'Trait' Analysis." *Journal of Experimental Education* 27: 107-24; December 1958. (b)

COOK, DESMOND L.; LINDEN, JAMES D.; and MCKAY, HARRISON E. "A Factor Analysis of Teacher Trainee Responses to Selected Personality Inventories." *Educational and Psychological Measurement* 21: 865-72; Winter 1961.

COSGROVE, DON J. "Diagnostic Rating of Teacher Performance." *Journal of Educational Psychology* 50: 200-204; October 1959.

COTSONAS, NICHOLAS J., JR., and KAISER, HENRY F. "A Factor Analysis of Students' and Administrators' Ratings of Clinical Teachers in a Medical School." *Journal of Educational Psychology* 53: 219-23; October 1962.

FLANDERS, NED A. *Interaction Analysis in the Classroom: A Manual for Observers.* Minneapolis: the Author (College of Education, University of Minnesota), 1960. 35 pp. (Mimeo.) (a)

FLANDERS, NED A. *Teacher Influence, Pupil Attitudes, and Achievement.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 397. Minneapolis: University of Minnesota, 1960. 245 pp. (b)

FLANDERS, NED A. "Analyzing Teacher Behavior." *Educational Leadership* 19: 173-75, 178-80, 200; December 1961.

GAGE, N. L., editor. *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. 1218 pp. (a)

GAGE, N. L. "Paradigms for Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 3, pp. 94-141. (b)

GAGE, N. L.; RUNKEL, PHILIP J.; and CHATTERJEE, BISHWA B. *Equilibrium Theory and Behavior Change: An Experiment in Feedback from Pupils to Teachers.* Urbana: Bureau of Educational Research, University of Illinois, 1960. 129 pp.

GETZELS, J. W., and JACKSON, P. W. "The Teacher's Personality and Characteristics." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 11, pp. 506-82.

HEIL, LOUIS M., and WASHBURNE, CARLETON. "Brooklyn College Research in Teacher Effectiveness." *Journal of Educational Research* 55: 347-51; May 1962.

HODGSON, THOMAS F., and HORST, PAUL. "The General and Primary Factors in Student Evaluation of Teaching Ability." (Abstract) *American Psychologist* 14: 374; July 1959.

HOYT, CYRIL J., and COOK, WALTER W. "The Stability of MTAI Scores During Two to Seven Years of Teaching." *Journal of Teacher Education* 11: 487-91; December 1960.

JACKSON, PHILIP W., issue editor. "Teacher Personality." *Theory into Practice* 2: 57-112; April 1963.

KEISLAR, EVAN R., and MCNEIL, JOHN D. "The Use of Pupil Accomplices To Investigate Teacher Behavior." *Journal of Experimental Education* 27: 237-40; March 1959.

KERLINGER, FRED N., and KAYA, ESIN. "The Predictive Validity of Scales Constructed To Measure Attitudes Toward Education." *Educational and Psychological Measurement* 19: 305-17; Autumn 1959.

KOUNIN, JACOB S., and GUMP, PAUL V. "The Comparative Influence of Punitive and Nonpunitive Teachers upon Children's Concepts of School Misconduct." *Journal of Educational Psychology* 52: 44-49; February 1961.

KOUNIN, JACOB S.; GUMP, PAUL V.; and RYAN, JAMES J., III. "Explorations in Classroom Management." *Journal of Teacher Education* 12: 235-46; June 1961.

LUCIO, WILLIAM H., and WENGER, M. A. *Prediction of Teacher Performance and Emotional Stability: A Psychophysiological Pilot Study of Female Student Teachers.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 648. Los Angeles: University of California, 1961. 91 pp.

Maccia, Elizabeth Steiner, and Maccia, George S. *The Way of Educational Theorizing Through Models.* Publication 62-111. Columbus: Bureau of Educational Research and Service, Ohio State University, 1962. 16 pp. (Mimeo.)

McKeachie, Wilbert J. "Research on Teaching at the College and University Level." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 23, pp. 1118-72.

Medley, Donald M. *The Development of Classroom Behavior Dimensions for Student Teachers and Pupils.* Paper presented to the American Association for the Advancement of Science, December 1962. New York: Division of Teacher Education, City University of New York, 1962. 17 pp. (Mimeo.)

Medley, Donald M., and Mitzel, Harold E. "Application of Analysis of Variance to the Estimation of the Reliability of Observations of Teachers' Classroom Behavior." *Journal of Experimental Education* 27: 23-35; September 1958. (a)

Medley, Donald M., and Mitzel, Harold E. "A Technique for Measuring Classroom Behavior." *Journal of Educational Psychology* 49: 86-92; April 1958. (b)

Medley, Donald M., and Mitzel, Harold E. "Some Behavioral Correlates of Teacher Effectiveness." *Journal of Educational Psychology* 50: 239-46; December 1959.

Medley, Donald M., and Mitzel, Harold E. "A Tentative Framework for the Study of Effective Teacher Behavior." *Journal of Experimental Education* 30: 317-20; June 1962.

Medley, Donald M., and Mitzel, Harold E. "Measuring Classroom Behavior by Systematic Observation." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 6, pp. 247-328.

Mitzel, Harold E. "Criteria of Teacher Effectiveness." *Encyclopedia of Educational Research.* Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1481-86.

National Education Association, Research Division. *Classroom Observation Record.* Research Memo 1962-31. Washington, D.C.: the Association, 1962. 16 pp. (a)

National Education Association, Research Division. "Profile of the Metropolitan Teacher." *NEA Research Bulletin* 40: 67-74; October 1962. (b)

National Education Association, Research Division. "What Do Teachers Think?" *NEA Research Bulletin* 40: 120-25; December 1962. (c)

National Education Association, Research Division. "Teachers in Public Schools." *NEA Research Bulletin* 41: 23-26; February 1963.

Nunnally, Jum C.; Thistlethwaite, Donald L.; and Wolfe, Sharon. "Factored Scales for Measuring Characteristics of College Environments." *Educational and Psychological Measurement* 23: 239-48; Summer 1963.

Page, Ellis B. "Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation." *Journal of Educational Psychology* 49: 173-81; August 1958.

Remmers, H. H. "Rating Methods in Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 7, pp. 329-78.

Rosencranz, Howard A., and Biddle, Bruce J. *The Role Approach to Teacher Competence.* Columbia: Social Psychology Laboratory, University of Missouri, 1962. 42 pp. (Mimeo.)

Rosenfeld, Howard, and Zander, Alvin. "The Influence of Teachers on Aspirations of Students." *Journal of Educational Psychology* 52: 1-11; February 1961.

Ryans, David G. *Characteristics of Teachers.* Washington, D.C.: American Council on Education, 1960. 416 pp. (a)

Ryans, David G. "Prediction of Teacher Effectiveness." *Encyclopedia of Educational Research.* Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1486-91. (b)

Ryans, David G. "Some Concomitants of Inventory-Estimated Emotional Maturity of Experienced Teachers." *California Journal of Educational Research* 11: 106-109, 144; May 1960. (c)

Ryans, David G. "Inventory Estimated Teacher Characteristics as Covariants of Observer Assessed Pupil Behavior." *Journal of Educational Psychology* 52: 91-97; April 1961. (a)

Ryans, David G. "Some Relationships Between Pupil Behavior and Certain Teacher Characteristics." *Journal of Educational Psychology* 52: 82-90; April 1961. (b)

Ryans, David G. *Application of Programmed Instruction and Auto-Instructional Devices in College and Their Relation to a Theory of Instruction.* SP-1084-000-01. Santa Monica, Calif.: System Development Corporation, 1963. 61 pp. (a)

Ryans, David G. *An Information System Approach to Theory of Instruction with Special Reference to the Teacher.* SP-1079. Santa Monica, Calif.: System Development Corporation, 1963. 65 pp. (b)

Ryans, David G. "Possible Directions for Teacher-Behavior Research." *Theory into Practice* 2: 105-12; April 1963. (c)

Schmid, John, and others. *An Analysis of Some Necessary Qualities of Teachers.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 453. Fayetteville: University of Arkansas, 1962. 101 pp.

Smith, B. Othanel. "A Concept of Teaching." *Teachers College Record* 61: 229-41; February 1960.

Smith, B. Othanel, and Meux, Milton O. *A Study of the Logic of Teaching.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 258. Urbana: Bureau of Educational Research, University of Illinois, 1962. 231 pp.

Smith, Louis M. *Pupil Expectations of Teacher Leadership Behavior.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 570. St. Louis: Graduate Institute of Education, Washington University, 1960. 35 pp.

Soar, Robert S. *Multivariate Statistical Procedures in Predicting Teacher-Pupil Classroom Behavior.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1170. Columbia: University of South Carolina, 1962. 167 pp.

Stern, George G. "Measuring Noncognitive Variables in Research on Teaching." *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. Chapter 9, pp. 398-447.

Stern, George G., and others. "Two Scales for the Assessment of Unconscious Motivations for Teaching." *Educational and Psychological Measurement* 20: 9-29; Spring 1960.

Torrance, E. Paul. *A Preliminary Analysis of the Evaluative Thinking of Effective and Ineffective Teachers of Experimental Mathematics Courses.* Minneapolis: Bureau of Educational Research, University of Minnesota, 1960. 12 pp.

Travers, Robert M. W., and others. *Measured Needs of Teachers and Their Behavior in the Classroom.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 444. Salt Lake City: Department of Educational Psychology, University of Utah, 1961. 200 pp.

Turner, Richard L., and Fattu, Nicholas A. *Problem Solving Proficiency Among Elementary School Teachers: I. The Development of Criteria.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 419. Bloomington: Institute of Educational Research, Indiana University, 1960. 60 pp.

Turner, Richard L., and Fattu, Nicholas A. "Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research." *Bulletin of the School of Education* (Indiana University) 36: 1-40; May 1960. (b)

Turner, Richard L., and Fattu, Nicholas A. "Skill in Teaching, Assessed on the Criterion of Problem-Solving." *Bulletin of the School of Education* (Indiana University) 37: 1-32; May 1961.

Turner, Richard L., and others. "Skill in Teaching, Assessed on the Criterion of Problem-Solving: Three Studies." *Bulletin of the School of Education* (Indiana University) 39: 1-31; January 1963.

Wallen, Norman E., and Travers, Robert M. W. "Analysis and Investigation of Teaching Methods." *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. Chapter 10, pp. 448-505.

Wallen, Norman E., and others. "Relationships Between Teacher Needs and Teacher Behavior in the Classroom." *Journal of Educational Psychology* 54: 23-32; February 1963.

Wilk, Roger E., and Edson, William H. *A Study of the Relationships Between Observed Classroom Behaviors of Elementary Student Teachers, Predictors of Those Behaviors, and Ratings by Supervisors.* Minneapolis: College of Education, University of Minnesota, 1962. 78 pp.

WITTROCK, M. C. "Set Applied to Student Teaching." *Journal of Educational Psychology* 53: 175-80; August 1962.

Additional References

ANDERSON, EARL W., and RUSHER, ELFREDA M. "Staff—Characteristics." *Encyclopedia of Educational Research*, Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1357-62.

BURKARD, M. INNOCENTIA. "Discernment of Teacher Characteristics by TAT Sequence Analysis." *Journal of Educational Psychology* 53: 279-87; December 1962.

COLE, DAVID L. "The Prediction of Teaching Performance." *Journal of Educational Research* 54: 345-48; May 1961.

GOWAN, J. C.; CONNER, CARITA; and KENNEDY, PHYLLIS. "A Follow-Up Study of Teaching Candidates." *Journal of Educational Research* 54: 353-55; May 1961.

HEIL, LOUIS M. *Modifying Behaviors (Self-Concept) of Certain Prospective Teachers.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1153. New York: Office of Testing and Research, Brooklyn College, 1962. 225 pp.

HUGHES, MARIE M., and OTHERS. *Development of the Means for the Assessment of the Quality of Teaching in Elementary Schools.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 353. Salt Lake City: University of Utah, 1959. 392 pp.

LANG, GERHARD. "Motives in Selecting Elementary and Secondary School Teaching." *Journal of Experimental Education* 29: 101-104; September 1960.

POPHAM, W. JAMES, and TRIMBLE, ROBERT R. "The Minnesota Teacher Attitude Inventory as an Index of General Teaching Competence." *Educational and Psychological Measurement* 20: 509-12; Autumn 1960.

SHELDON, M. STEPHEN; COALE, JACK M.; and COPPLE, ROCKNE. "Concurrent Validity of the 'Warm Teacher Scales.'" *Journal of Educational Psychology* 50: 37-40; February 1959.

SYMONDS, PERCIVAL M. "An Educational Interest Inventory." *Educational and Psychological Measurement* 18: 377-85; Summer 1958.

WILSON, W. CODY, and GOETHALS, GEORGE W. "The Relation Between Teachers' Backgrounds and Their Educational Values." *Journal of Educational Psychology* 51: 292-98; October 1960.

CHAPTER VIII

Teacher Personnel: A Brief Evaluation of the Research Reviewed

WILLIAM B. MICHAEL

In his recent evaluation of research reviewed in each of the chapters of the June 1963 Review, Alexander posed five different questions, the answers to which furnished a basis for the formulation of judgments concerning both the worth of the research reported and the direction of future research. Slightly modified, these same five questions may be used as standards against which to evaluate the research pertaining to teacher personnel: Which studies reviewed seem to be critically significant? What new and promising areas of research concerned with problems of teacher personnel have been opened? What significant contributions to a theoretical conceptualization of problem areas in teacher personnel have been proposed? What new and promising methodological approaches to research concerning teacher personnel have been introduced? What lies ahead in research for teacher personnel?

To a great extent, the answers to these questions depend on the manner in which the reader defines the term *research* in education. If one sets his sights rather high, it seems that studies of a somewhat novel and imaginative nature that one can relate to a theoretical framework from which testable hypotheses are deduced can qualify as significant research, especially if the design and subsequent statistical treatment are both sound and sophisticated. Many research workers believe that mere status studies in which large amounts of demographic and other descriptive data are systematically gathered and enumerated in endless tables and charts do not constitute research. Although such studies may contribute little to fundamental knowledge, they frequently serve to answer practical questions of immediate urgency and to yield a basis for prediction of trends. In particular, investigations of the survey type are useful in determination of requirements relating to recruitment, promotion, and retention of personnel; to physical facilities; to school finance and tax rates; and to countless operational problems. There is also a place in educational research for expository articles that represent a scholarly attempt to analyze, to synthesize, to unify, and to interrelate important research findings. Furthermore, papers concerned with philosophical and historical issues, with problems about value judgments, and with methodological and statistical developments can represent significant contributions to educational research.

To a considerable extent, however, it appears that those research papers that are related to a theoretical framework constitute the ones most likely to contain the findings necessary for the development of a science of education. If one examines recent issues of the REVIEW and compares them with those of five or ten years ago, it appears that there is an increase in the proportion of articles cited in the bibliographies that reflect concern with experimental verification of hypotheses or of predictions derived from a theoretical structure. Such articles should be given high priority for inclusion in the REVIEW. In this issue, the reader will note that except for Chapter VII relatively little of the research reported can be regarded as theoretical in its orientation—a fact which is due probably in part to the nature of the topics considered.

High Standards Needed in Selecting Publications for the REVIEW

Irrespective of what types of research studies are to be described in future issues of the REVIEW, it is the writer's contention that authors should give greater attention than they have in the past to establishing a set of relatively rigorous standards by which they can decide whether to include a given article, report, or book. The rate of publication is increasing so rapidly that only a small fraction of the total number of articles in the professional literature can be cited. Probably, on the average, only 10 percent of published papers in educational journals are worthy of being reported in the REVIEW.

The writer's bias was shared by a group of contributors to Chapter V of the June 1963 REVIEW who established and applied specific standards to the selection of a relatively small number of references from a relatively large tentative bibliography. Such criteria as the relevance of the problem, adequacy of the experimental design, and appropriateness of statistical treatment had to be met. In a pointed statement that constituted their closing paragraph, Gibboney and others wrote:

> "This journal could perform a vital service by rigorously selecting fewer studies for review, by analyzing each study in depth, and by candidly reporting its strengths and weaknesses. Exclusion from the REVIEW would increasingly become the vehicle for rigorous internal professional criticism."

Published Works of Critical Significance

In his conscientious efforts to be highly selective, each contributor to this issue of the REVIEW not only tried to include significant studies but also endeavored to cover each of the major problem areas within his assigned topic. Several of the contributors who indicated marked conflicts concerning the inclusion of certain studies of questionable quality decided that they had an obligation to report a certain amount of research in each topical area, even at the risk of including the most promising studies among several that seemed to constitute mediocre and trivial research. It may well be that certain topics should have been omitted. In the writer's estimation, slightly more than 55 percent of the entries in the bibliographies could have been deleted without any noticeable depreciation in the value of the current issue. This opinion is based on a careful examination of the contents of approximately 87 percent of the references cited.

Without doubt, the most significant contribution during the past five years was the AERA-sponsored *Handbook of Research on Teaching*, edited by Gage (1963a). The strong emphases on both a theoretical framework and research methodology, the attempts at a description of the major variables and areas of research on teaching, and the detailed exposition of research on teaching at various grade levels and in different subject matter areas should serve to encourage its readers to raise the level of their own future research endeavors that are concerned with teaching behavior, the instructional process, and the related areas of assessment, selection, and recruitment. Careful reading of chapters of this volume not only will suggest numerous research problems worthy of investigation but also will point to improvements both in design of studies and in the treatment of statistical data. One may hope that in time the stimulation which this volume affords will bring about a body of knowledge concerning teaching and instruction that will have a direct impact on the problems of teacher certification, the nature of pre-service and in-service education of teachers, the conditions of employment and service in elementary and secondary schools as well as in higher education, and the realization of improvements in the economic, legal, and social status of teachers.

Another noteworthy contribution was Ryans' (1960) *Characteristics of Teachers*, which brought together the extensive results of a large-scale research program concerning the assessment of teacher behavior. The contents of this monumental volume have been summarized briefly by Ryans in Chapter VII of this issue.

Although most of the significant contributions and best-designed investigations have appeared in conjunction with the conceptualization and assessment of both the properties of teacher behavior and the characteristics of the instructional process, the writer would like to call attention to two comprehensive research studies conducted by the National Education Association, Research Division (1960, 1962) concerning teacher supply and demand. On the basis of findings of these two studies as well as the material in his own book (Stone, 1961), the author of Chapter I concluded that an important trend exists relative to teacher supply and demand—namely, that with respect to the quality of teachers recruited the gap is widening between those states with relatively high standards and those states with relatively low standards. Closely related to the fact that those states with higher requirements tend to obtain the services of the better qualified teachers was Reynard's observation—from numerous studies reviewed in Chapter III—that the preparation of teachers is being improved and extended in general education, in the field of academic and technical specialization, and in professional education.

The conclusions that several of the contributors discerningly reached from their reviews of the literature were indeed disappointing. In Chapter II, Durflinger decried the lack of replication of investigations and the lack of any unified or substantial cooperative efforts in the large number of reported empirical studies that were concerned with recruitment and selection of teachers. Even less encouraging than Durflinger's conclusion was Gordon's pointed criticism in Chapter IV of the barren and unimaginatively repetitive nature of research regarding conditions of employment and service in elementary and secondary schools—an emptiness that was attributed to lack of a clear definition of the role of education and of the teacher in American society. In a similar vein, Marks and Dillon noted in Chapter V a marked absence of systematic research about problems related to teacher personnel in higher education. Although sterility in concept also was apparent in those phases of Chapter VI concerned with economic and legal status of teachers, some measure of hope could be taken from those cited studies regarding the social status of teachers in which role perceptions and image theory afforded a number of promising hypotheses worthy of additional investigation.

New and Promising Areas of Research

From the reviews of the published literature during the past five years, it is with respect to the teaching process itself that the greatest potentiality for research lies, as inspection of the *Handbook of Research on Teaching* will readily reveal. In particular, the paradigms for research on teaching reported by Gage (1963b) hold great promise for guiding future research. Similarly, the numerous models which Ryans described in Chapter VII under the heading of Theory of Instruction and Teacher Behavior constitute important steps forward both in the development of conceptual frameworks within which the teaching-learning process may be viewed and in the heuristic values which these models possess.

Contributions to Theory

As indicated in the previous paragraph, the construction of comprehensive paradigms and models related to the instructional and learning process has been an important development—perhaps even a major breakthrough. It appears that the educational psychologists have moved quickly in directions that their colleagues in experimental and theoretical psychology in Division 3 of the American Psychological Association (APA) have been following for several years. A most encouraging sign has been that the educational psychologists who contributed important chapters to the *Handbook of Research on Teaching* not only have manifested a marked degree of sophistication about model building but also have given ample evidence of their concern with the applications afforded by theoretical conceptualizations to problems of teaching in the classroom. It is encouraging to see that all of the highly talented psychologists who are accomplished in theory building do not reside in the exclusive community of APA's Division 3.

New and Promising Procedures of Research

As mentioned previously, it is in the study of teacher behavior that innovations in research methodology have been most prominent. In Chapter VII, specific attention was given to the criterion problem. Six chapters in the *Handbook of Research on Teaching* were written concerning methodologies of research in teaching. For the whole area of educational research that is quantitatively oriented, these chapters will represent keystones which research workers will wish to consult.

Future of Research in Teacher Personnel

As of the moment, the recent theoretical advances concerning the assessment of teacher behavior and instruction hold tremendous promise for numerous imaginative and creative contributions to a science of education during the next several years. The theoretical models proposed should serve as important aids to the conceptualization of the criterion problem. Along with improvements and refinements in observational, statistical, and data-processing techniques, the future seems brighter than it has ever been. Whether the data from forthcoming experimental studies will serve to validate the constructs underlying the theoretical formulations remains to be seen. If the results of carefully designed experiments bring about revisions in the multivariate models that will make them workable in the classroom, then important advances can be expected not only in the teaching process itself but also in the development of instruments of selection, in the procedures of recruitment, and in the training of teachers. Once the criterion can be adequately described and replicated, it should be relatively easy to develop reliable and valid predictors. It may well be that a turning point has been reached in research on the teaching process which will be reflected in substantial gains in the understanding of its nature.

In most of the other problem areas of teacher personnel, it may be expected that large numbers of survey-research studies will continue to be made, although it can be hoped that theoretical considerations will eventually assume some degree of importance. A fundamental weakness of many investigations concerned with legal, economic, and related demographic problems has been neglect of what the aims and purposes of American education really are. Value-oriented hypotheses formulated with the aid of educational philosophers and historians, psychologists, sociologists, and anthropologists could do much to furnish a social framework within which status studies could be made increasingly meaningful.

In view of the uneven rate of growth of the research contributions from each of the conventionally viewed areas of teacher personnel, it may be advisable in the future that two types of issues of the REVIEW be developed. First, problems of the assessment of teacher behavior, teacher selection, teacher training, and possibly social status might be in one cycle of publication. On the other hand, separate issues of the REVIEW could be devoted at irregular intervals to administrative problems concerned with recruitment; with certification, supply, and demand; and with economic, legal, and related demographic characteristics. At the moment, it appears that the significance of educational research is somewhat greater in the realm of teacher behavior and the instructional process than it is in other topics of teacher personnel. Perhaps issues every three years and every six years, respectively, would be advisable for reporting (a) research on the problems directly related to teaching behavior itself and to the instructional process and (b) studies on the relatively mechanistic administrative problems underlying the professional activities of teachers.

Such a separation, however, may introduce certain risks. Indeed, the opportunity for transfer effects of research efforts concerning the process of teaching to related administrative procedures would be curtailed. A vague, ill-defined unity of topical areas in teacher personnel, which may be better than no unity at all, could be destroyed. Nevertheless, if the level of sophistication in research concerned with teacher behavior continues to grow as rapidly as it has during the past five years and if the research concerned with other aspects of teacher personnel continues to follow its same unimaginative, repetitive pattern, the differences in the types of audiences interested in these two broad areas of teacher personnel can be expected to become increasingly divergent. Within the next two or three years, the answer may be forthcoming concerning whether the psychology and sociology of the teaching process has indeed achieved a position of eminence apart from other problem areas underlying the global term *teacher personnel.*

Bibliography

GAGE, N. L., editor. *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. 1218 pp. (a)

GAGE, N. L. 'Paradigms for Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 3, pp. 94-141. (b)

NATIONAL EDUCATION ASSOCIATION, RESEARCH DIVISION. *Teacher Supply and Demand in Public Schools, 1960.* Research Report 1960-R7. Washington, D.C.: the Association, 1960. 48 pp.

NATIONAL EDUCATION ASSOCIATION, RESEARCH DIVISION. *Teacher Supply and Demand in Public Schools, 1962.* Research Report 1962-R8. Washington, D.C.: the Association, 1962. 55 pp.

RYANS, DAVID G. *Characteristics of Teachers.* Washington, D.C.: American Council on Education, 1960. 416 pp.

STONE, JAMES C. *California's Commitment to Public Education.* New York: Thomas Y. Crowell Co., 1961. 193 pp.

INDEX TO VOLUME XXXIII, NO. 4

Page citations are made to single pages; these are often the beginning of a chapter, section, or running discussion dealing with a topic.

Assessment of teacher behavior and instruction: criterion problem, 422; future research, 433; general references concerning, 416; information system model, research literature viewed within framework of, 420; interpersonal relations in classroom, 424; limitations in findings, 431; methodology, 422; observation, 423; overview of trends in research, 415; scheme for classifying teacher behaviors, 423; teacher effectiveness, 431; theory of instruction and teacher behavior, 419

Autonomy of teachers: 402

Campus laboratory school: 377

Certification: levels of responsibilities, 344; state-by-state variations, 343; teacher examinations, 345; trends, 346; variations in, 344

College instruction: research in, 431

Conditions of employment and service in elementary and secondary schools: economic status and economic conditions of employment, 386; employment, assignment, and turnover, 385; evaluation, retention, and dismissal, 385; needed research in, 388; staff organization and utilization, 383; teacher load, 384; teacher morale and job satisfaction, 387; teacher status and role expectations, 382

Contributions to theory: 445

Criterion problem: in assessment of teaching, 422

Description of teacher behavior based on: direct observation, 428; inferences, 428

Dismissal of teachers: 385

Economic status and economic conditions of employment for teachers: 386

Elementary school teacher education: 370

Employment, assignment, and turnover: 385

Evaluation: in elementary and secondary schools, 385; in higher education, 392; of research reviewed in issue, 442

Factor studies: contributions to identification of characteristics in training and selection of teachers, 364

Future of research in teacher personnel: 446

High standards needed in selecting publications for the REVIEW: 443

Information system model of teaching: external inputs—conditions external to the teacher, 426; internal inputs contributing to instructor behavior, 424; products of instructor behaviors—pupil behaviors, 429; teacher behaviors—outputs of teacher information processing, 427; teacher information processing behavior, 426

In-service education: of teachers, 373

Instructor, or teacher, behavior: teacher-student variables in, 424

Legal status of teachers: certification, tenure, and autonomy, pertaining to, 402; civil rights, regarding, 403; compensation, residual rights, and royalties in relation to television, radio, and programed instruction, 404; retirement, with reference to, 405; teacher-board relations, concerning, 404; teacher-pupil relations, concerning, 404

Load for teachers: 384

Methodology in assessment of teacher behavior and instruction: criterion problem, 422; interpersonal relations in the classroom, 424; observation processes, 423; scheme for classifying teacher behaviors, 423

Morale, opinions, and attitudes: of teachers, 409

Need and preparation of teachers: in higher education, 391
New and promising areas of research: in teacher personnel, 445
New and promising procedures of research: in teacher personnel, 446
New media: in teacher education, 374

Perception-diagnosis theory of teaching: 419
Personality characteristics: in teacher selection, 361
Pre-service programs: 370
Pre-service and in-service education for teachers: campus laboratory school, role of, 377; elementary teachers, concerning, 370; in-service programs, 373; new media, with reference to, 374; needed research in, 378; professional laboratory experiences, concerning, 375; secondary teachers, concerning, 371
Published works: of critical significance in area of teacher personnel, 443
Pupil behaviors: products of instructor behaviors, 429; pupil achievement and learning, 430

Recruitment: age at making decision to be teacher, 356; factors in decision making, 356
Recruitment and selection of prospective elementary and secondary school teachers: factor studies, 364; recruitment, aspects of, 355; selection, factors pertaining to, 358
Research needed concerning: conditions of employment and service in elementary and secondary schools, 388; pre-service and in-service teacher education, 378; teacher certification, supply, and demand, 351; understanding teacher behavior, 433
Retention of teachers: 385

Salaries and other benefits: in elementary and secondary education, 386; in higher education, 394; for teachers in general, 398; for women teachers, 399; in relation to additional employment, 400
Scheme for classifying teacher behaviors: 423
Secondary school teacher education: 371
Selection, factors pertaining to: academic competence, 359; career commitment, 361; criteria, 358; personality characteristics, 361
Social status of teachers in relation to: educational level, 406; images and stereotypes of teachers and teaching, 407; occupation of father, 407
Staff organization and utilization: 383
Summary of teacher-student variables: as considered in the literature, 424
Supply and demand: economic factors, 347; needed research in, 351; remedies for the shortage, 350; subject and grade level factors, 348; other factors, 349

Teacher as problem solver: model for, 421
Teacher attitudes and opinions: 409
Teacher-board relations: legal aspects of, 404
Teacher certification, supply, and demand: 343
Teacher education: 369
Teacher effectiveness: 431
Teacher examinations: 345
Teacher load: 384
Teacher morale: 387
Teacher personnel research in higher education: need and preparation of teachers, 391; new data sources concerning, 396; ratings and evaluation in, 392; salaries and other benefits, pertaining to, 394
Teacher-pupil relations: legal aspects of, 404
Teacher status and role expectations: 382
Teacher-student variables in relation to systems approach: 424
Teaching as system of social action: 419
Tenure of teachers: 402
Theory of instruction and teacher behavior: contributions in general, 445; information system model, 420; model viewing teacher as problem solver, 421; perception-diagnosis theory of teaching, 419; teaching as system of social action, 419

REVIEW OF EDUCATIONAL RESEARCH

Official Publication of the American Educational Research Association
Contents are listed in the Education Index.

Copyright 1963
By National Education Association of the United States, Washington, D.C.

Volume XXXIII, No. 5 **December 1963**

Statistical Methodology

Reviews the literature for the three-year period
since the issuance of Vol. XXX, No. 5, December 1960.

TABLE OF CONTENTS

Chapter *Page*

Foreword 454

I. Hypothesis Testing in Relation to Statistical Methodology.... 455
CHERRY ANN CLARK, *The Meyers Clinic, Los Angeles, and Claremont Graduate School, Claremont, California*

II. Selected Contributions to Parametric and Nonparametric Statistics 474
WILLIAM B. MICHAEL, *University of California, Santa Barbara, California*

III. Developments in Analysis of Variance and Design of Experiments 490
DEE W. NORTON, *University of Iowa, Iowa City, Iowa*

IV. Regression and Correlation.. 501
WILLIAM J. MOONAN, *U.S. Naval Personnel Research Activity, San Diego, California*
JOHN H. WOLFE, *U.S. Naval Personnel Research Activity, San Diego, California*

Chapter *Page*

V. Factor Analytic Methodology.......................... 510

Philip R. Merrifield, *System Development Corporation, Santa Monica, California*

Norman Cliff, *University of Southern California, Los Angeles, California*

VI. Design of Sample Surveys in Education.................. 523

Francis G. Cornell, *Educational Research Services, Inc., White Plains, New York*

Eugene P. McLoone, *Office of Education, U.S. Department of Health, Education, and Welfare, Washington, D.C.*

VII. Scaling and Measurement Theory........................ 533

Harold Schiffman, *Educational Testing Service, Princeton, New Jersey*

Samuel Messick, *Educational Testing Service, Princeton, New Jersey*

VIII. Statistical Learning Models.......................... 543

Ronald A. Weitzman, *Los Angeles State College, Los Angeles, California*

IX. Decision Theory and Utility Management................. 556

Robert W. Earl, *Claremont Graduate School, Claremont, California*

X. Use of Computers in Educational Research................ 566

Frank B. Baker, *University of Wisconsin, Madison, Wisconsin*

Index 579

This issue of the REVIEW was prepared by the
Committee on Statistical Methodology

William B. Michael, *Chairman*, University of California, Santa Barbara, California
Raymond O. Collier, University of Minnesota, Minneapolis, Minnesota
Henry F. Kaiser, University of Illinois, Urbana, Illinois
Julian Stanley, University of Wisconsin, Madison, Wisconsin

with the assistance of

Frank B. Baker, University of Wisconsin, Madison, Wisconsin
Cherry Ann Clark, The Meyers Clinic, Los Angeles, and Claremont Graduate School, Claremont, California
Norman Cliff, University of Southern California, Los Angeles, California
Francis G. Cornell, Educational Research Services, Inc., White Plains, New York
Robert W. Earl, Claremont Graduate School, Claremont, California
Eugene P. McLoone, Office of Education, U.S. Department of Health, Education, and Welfare, Washington, D.C.
Philip R. Merrifield, System Development Corporation, Santa Monica, California
Samuel Messick, Educational Testing Service, Princeton, New Jersey
William J. Moonan, U.S. Naval Personnel Research Activity, San Diego, California
Dee W. Norton, University of Iowa, Iowa City, Iowa
Harold Schiffman, Educational Testing Service, Princeton, New Jersey
Ronald A. Weitzman, Los Angeles State College, Los Angeles, California
John H. Wolfe, U.S. Naval Personnel Research Activity, San Diego, California

FOREWORD

It has been nine years since there has been an issue of the REVIEW devoted exclusively to statistical methodology. The rate at which publication is increasing is nothing short of astounding. Likewise, the level of mathematical maturity required of the reader of statistical literature is rising at a correspondingly rapid rate. One dare not think what the status of statistical methodology will be by the year 1972!

Despite the almost formidable advances in statistical knowledge, the author or authors of each chapter have made a sincere attempt to capture the conceptual significance of the methodological developments with which they have been dealing. To translate concepts from a largely mathematical formulation to a verbal medium inevitably introduces errors in communication—especially in the direction of oversimplification. Since most readers of the REVIEW have not specialized in statistics, the attempt to convey in verbal terms the major developments in statistical methodology seems to be a desirable and defensible objective. How well the writers of the current issue have succeeded must await the reaction of the readers.

It should be noted that since chapters on statistical methodology, experimental design, scaling theory, and data processing appeared in the December 1957 and the December 1960 issues of the REVIEW, not all chapters in the present issue furnish coverage for a nine-year period. Chapters I, VIII, and IX afford a review of contributions during the nine-year period from 1954 to 1963, although relatively greater emphasis has been placed on the developments of the past three years in contrast to that put on those contributions during the prior six years, 1954-1960.

The Chairman would like to express his appreciation not only to the other members of the Committee and the authors but also to the important individuals behind the scenes who make possible the production of each issue of the REVIEW. In particular, he wishes to thank Mrs. Hazel Blankenchip who typed certain chapters and retyped portions of others that had been edited. She has made an important contribution to the preparation of each issue of the REVIEW since April 1962. Deep appreciation is expressed to Linda Thurston Reeder for her painstaking efforts in technical editing, for her attention to numerous important points of form and style, and for her constructive suggestions in the preparation of the manuscripts.

Comments from readers concerning the contents of this issue and suggestions for future issues that are devoted to research and statistical methodology will be most welcome.

William B. Michael, *Chairman,*
Committee on Statistical Methodology

CHAPTER I

Hypothesis Testing in Relation to Statistical Methodology

CHERRY ANN CLARK

THE SHORTCOMINGS in the methodology of statistical hypothesis testing used in educational and psychological research have been emphasized repeatedly in recent behavioral science and statistical literature (Binder, 1963; Edwards, Lindman, and Savage, 1963; Grant, 1962; Lubin, 1962; McNemar, 1960; Mowrer, 1960; Nunnally, 1960; Rozeboom, 1960; Savage, 1957). This chapter reviews the salient points of the criticisms. Since this issue of the REVIEW marks the first time an entire chapter has been devoted to the statistical methodology of hypothesis testing, a brief account of several theories of statistical inference is included to provide a background for evaluating the rationales of significance tests compared with other methods for statistical inferences, as well as the modifications in the function of the null hypothesis in testing statistical hypotheses. The problems in statistical inference which have been associated with the widespread use of significance tests are reviewed. The limitations in and the effectiveness of significance tests as methods for informative inference are described. The applications of interval estimation contrasted with significance tests in the investigation of hypotheses and models and in the development of a body of empirical data are summarized.

Hypothesis Testing and Statistical Inference

Hypothesis testing is a central and complex problem in the methodology of science. It involves comparing the deductions or the predictions from scientific hypotheses with observational data to eliminate unsatisfactory hypotheses and to give support to satisfactory ones.

A primary objective of hypothesis testing methodology is the formulation of rules or criteria, sometimes called decision procedures, to use in determining whether the data should be construed as rejecting or accepting the hypothesis under investigation. Decision procedures are judged by such criteria as consistency, relevance, completeness, and effectiveness. Decision rules are formulated to provide objective, reliable, and valid solutions to important problems in the analysis of data (Buehler, 1959; Hotelling, 1958; Jeffreys, 1957, 1961; Tukey, 1962). The hypotheses to be tested must be examined for their logical validity or consistency, their heuristic value, and their amenability to empirical test. The testing of hypotheses requires that appropriate controls be employed in making observations and in analyzing the data. Systematic errors in the experimental procedures must be minimized if decision rules are to be efficient (Tukey, 1954, 1960a, 1962).

Statistical hypothesis testing is a special instance of the general scientific method of testing hypotheses. As in the general method, statistical hypothesis testing combines deductive and inductive methods in intricate ways. For the most part, statistical hypothesis testing has been concerned with determining which one of a dichotomous set of mutually exclusive and exhaustive hypotheses is to be rejected and which one accepted at a specified level of risk in making an erroneous conclusion or decision. Probability theory provides the deductive foundations for theories of statistical inference (Hotelling, 1958). Natural phenomena which are presumed to be subject to random fluctuations, either because of a stochastic process characterizing their behavior or because of random errors in observing or measuring them, provide the inductive basis for statistical inference. Both enumerative and eliminative methods of induction are used in treating the data and in formulating statistical generalizations.

The question of whether statistical methods based on eliminative or enumerative rules of induction constitute the best foundation for statistical inference is widely disputed (Rozeboom, 1960; Savage and others, 1962). The argument about whether scientific knowledge advances primarily by accumulation of positive instances of phenomena to support a hypothesis or by negative instances to refute a hypothesis is the focus of some of the criticism of the classical theories of statistical inference—namely, the Fisher and Neyman-Pearson theories (Savage, 1961). Another argument in the current controversy on the foundations of statistics focuses on the question of whether there is a sound logical basis for assigning a direct probability to a set of propositions or hypotheses. The frequency or objective school has denied the legitimacy of such a procedure. Probability, for this school, refers to the relative frequency of events within a class of random recurrences (Hotelling, 1958), not to the amount of support for or the degree of belief in a hypothesis. On the other hand, the subjective probability school, among others, has argued that a fruitful and consistent way to proceed with statistical inferences is by assigning a degree of belief to a statement based upon available information and then by altering the amount of uncertainty about the statement in the light of additional experimental evidence by the use of Bayes's theorem (Lindley, 1953, 1961; Rozeboom, 1960; Savage, 1961).

Another aspect of the controversy is how the likelihood function should be incorporated into statistical theory and practice. One of the serious limitations in the classical theories of significance tests is that they are not dependent upon the likelihood function (Birnbaum, 1962).

Statistical hypotheses are more restricted than are scientific hypotheses, just as statistical inferences are more restricted than scientific inferences (Bolles, 1962; Cox, 1958b; Tukey, 1960a). Statistical hypotheses concern the behavior of observable random variables, whereas scientific hypotheses treat the phenomena of nature and man. A null hypothesis is a particular statistical hypothesis which refers to the theoretical probability distribution governing the random variable(s) under investigation and is the hypothesis under test (Kendall and Stuart, 1961). Statistical inference, broadly defined, deals with statements about statistical populations made from given observations with a measured degree of uncertainty (Cox, 1958b). Statistical inference proceeds from observations to conclusions about the populations sampled. Scientific inference, on the other hand, often argues from descriptive facts about populations to an abstraction about the system of phenomena under investigation. A statistical inference frequently is instrumental in formulating a general inference, but usually it constitutes only a small part of the uncertainty connected with the scientific inference (Cox, 1958b; Tukey, 1960a).

Decision procedures used in making statistical inferences, including those that aid in the selection of statistical design and analysis (Cox, 1958a; Savage, 1961), require that the investigator use his judgment in diverse ways. How such judgment should be allowed to affect the use of the decision procedures is a crucial issue in the present controversy on the foundations of statistical inference (Savage and others, 1962; Tukey, 1954, 1960a, 1962). Much of the criticism of the rationales for significance tests has been aimed at the inconsistencies in the performance of the decision procedures (Berkson, 1942; Birnbaum, 1962; Cox, 1958b; Edwards, Lindman, and Savage, 1963; Pratt, 1961a, b; Savage and others, 1962).

Development of Theories of Statistical Inference and Significance Tests

The use of significance tests preceded the development of any systematic theories of statistical inference. Barnard (Savage and others, 1962) and Lehmann (1959) have mentioned their use by two early probability theorists, Daniel Bernoulli and Pierre Simon de Laplace. The principal components of significance tests were contained in their mathematical studies on the departures of planetary orbits from specific mathematical models. The essential components were the theoretical or mathematical hypotheses and equations and the observations on the extent of the departures of the planetary planes from the expected values. The significant departures were those that were considered unlikely on a particular hypothesis. When Bernoulli found that the values he computed were uniformly improbable on the hypotheses, he concluded that the observations provided evidence for rejecting the initial hypotheses. He did not consider an alternative hypothesis, nor did he have a particular degree of departure in mind before he made his computations.

Karl Pearson's work on goodness of fit tests and the chi square distribution and W. S. Gossett's derivation of the student distribution, a small sample exact-sampling distribution, were important influences on subsequent developments in the theory of significance tests.

Contributions of Fisher to Statistical Inference

Sir Ronald Fisher has been called the founder of modern statistical theory. He propounded many basic concepts and methods and developed the first complete theory of statistical inference (Hotelling, 1951; Pearson, 1962; Yates, 1951). Among his contributions to the theory of significance tests were a systematic rationale for critical ratio tests based on the Central Limit Theorem and the Law of Large Numbers; exact small sample tests for which he derived many theoretical sampling distributions; and randomization or permutation tests to be used in complex experimental designs when the assumptions underlying other significance tests procedures are not fulfilled. He developed interval estimation methods, which culminated in his theory of fiducial inference with its fiducial intervals. He introduced an extensive formal theory of experimental and statistical design and related methods of analysis, including the analyses of variance and covariance. He set forth the method of maximum likelihood as a foundation for statistical inference and as a basis for developing many statistical techniques. He discussed the concept of sufficient statistics, which has been used extensively in modern statistical theories. He argued that inverse probability has no place in modern science or statistical theory (Fisher, 1956, 1960).

He originated the term *null hypothesis.* He presented significance tests of null hypotheses as examples of a logical disjunction. In this view, the null hypothesis could never be proved by experimentation. On the contrary, the null hypothesis exists only to be rejected by a sufficiently sensitive experiment. Rarely is an experimenter interested in accepting the null hypothesis, an exception being in determining the uniformity of experimental procedures (Fisher, 1960). Fisher discussed the combination of estimation procedures with significance tests after the initial phases of experimentation have been completed.

Hotelling (1951) and Yates (1951), in evaluating Fisher's impact upon statistical practice in the social sciences, suggested that his books have not had an altogether positive influence. Psychologists, for example, have been content to publish very limited reports on one or at most several significance tests based on null hypotheses of no difference (Grant, 1962; Nunnally, 1960; Rozeboom, 1960). Such statistical tests have very little informative value. They often are known to be false prior to experimentation (Savage, 1954).

Contributions of Neyman and Pearson to Statistical Inference

Jerzy Neyman and Egon Pearson collaborated for more than a decade on the development of many aspects of statistical theory and methods.

Some of their work was an extension of Fisher's contributions, and some of it was a reaction to Fisher's proposals. They expanded and systematized the rationale of significance tests. They argued for the necessity of carefully considering the alternative to the null hypothesis in making any test of significance. They introduced the notions of errors of the first and second kinds, the power of a test in discriminating between the hypothesis under test and the alternative, and the comparison of the power functions of different tests to aid in the selection of the most satisfactory statistical test. They formulated an extensive rationale for the comparison and derivation of different tests in terms of their "nice properties," such as the uniformly most powerful test, most powerful tests, and unbiased tests. They expanded Fisher's maximum likelihood method into the test criterion called the Neyman-Pearson lemma, which they used to develop tests of simple statistical hypotheses. They gave considerable attention to the likelihood ratio in the development of likelihood ratio tests, including both univariate and multivariate tests. They pointed out the great complexity of the problems in deriving suitable test criteria for complex hypotheses as contrasted with simple hypotheses. They, like Fisher, demonstrated the wide variety of statistical tests to be derived from the maximum likelihood method.

They posited that the size of the sample should be set before experimentation is begun. The decision rules for tests of significance within their theory are based upon hypothetical repetitions of similar size random samples from the population(s) under investigation. Determination of the size of the sample is a function both of the size of the critical region and of the sensitivity or power of the test in detecting departures from the null hypothesis. They conceptualized the error rate in significance tests as the probability of obtaining a given or more extreme level of significance in an extended series of random fixed sample size experiments. They reasoned that a significance test is a decision procedure to be used as an aid in deciding which one of two actions to take in the face of uncertainty (Lehmann, 1959).

In their exposition of the theory of significance tests, Neyman and Pearson incorporated the spirit of probabilistic reasoning for a two-valued problem. Theoretically, the null hypothesis and the alternative hypothesis have equal status: the data are to determine which hypothesis is more probable. The error of the first kind is considered the more serious; it is determined by the size of the critical region or the level of significance. The error of the second kind cannot be controlled directly by the experimenter, but is a function of the size of the critical region and of the characteristics of the power function of the test in relation to the specific alternative hypothesis under consideration. In actuality, the null hypothesis is rarely found to be acceptable in the Neyman-Pearson formulation, for as the sample size increases, a sufficiently sensitive test rejects the null hypothesis at a decreasing level of significance (Kendall and Stuart, 1961; Pratt, 1961b). This fact has been of grave concern to a number of statisticians (Savage and others, 1962). This fact, in addition to certain other inconsistencies in the performance of the decisions rule, has evoked sharp criticism of this classical method for statistical inference (Pratt, 1961b).

The Neyman-Pearson theory indicated the relationship between tests of significance and the method of interval estimation called confidence intervals (Berkson, 1942; Birnbaum, 1961; Bulmer, 1957; Kendall and Stuart, 1961; Lehmann, 1959; Natrella, 1960; Pratt, 1961a). A test of significance is concerned with only those values specified by the hypotheses under consideration, while confidence intervals by their width and by the level of confidence indicate the multitude of statistical hypotheses which are acceptable as well as how acceptable each hypothesis is by its location within the confidence interval. The Neyman-Pearson rationale showed the practical and theoretical equivalence of tests of significance accompanied by the power function or the operating characteristic curve and confidence intervals (Natrella, 1960). Pratt (1961a, b) reminded statisticians of one important shortcoming of the theory of confidence intervals: the probability is either one or zero that the value of interest is contained within the confidence interval; moreover, the formulation is based upon the frequency interpretation of probability, which does not allow for a straightforward degree of confidence in the likelihood of the obtained statistic.

Educational and psychological investigators have rarely used effectively the Neyman-Pearson formulation either for significance testing or for interval estimation.

Contributions of Wald to Statistical Inference

In the short but exceptionally productive time that Wald contributed to statistics, he made many important modifications in the theory of statistical hypothesis testing. He recognized the desirability of incorporating into statistical theory and practice provision for handling sequentially selected samples of varying sizes of items rather than limiting statistical design and analysis to fixed sample sizes, as required by the Neyman-Pearson method. He devised the sequential probability ratio test to analyze sequentially sampled items in industrial situations, such as quality control. Sampling is continued until a desired degree of precision is obtained as a basis either for accepting the null hypothesis or for rejecting it and accepting the alternative. The dichotomous decision problem of earlier significance tests is expanded into a trichotomous decision problem: either to accept or reject a statistical hypothesis or to continue sampling (Savage, 1954; Schlaifer, 1959). Two constants are selected for the test procedure to give the desired weight to the probability of each of the two kinds of errors.

In his *Statistical Decision Functions*, Wald (1950) gave explicit consideration to the statistician's role as a decision maker by his detailed specifications of the various factors of the decision situation. Previously, Neyman and Pearson had given indirect consideration to the possible losses consequent to a wrong decision by the different values to be selected for the two kinds of errors. Instead of basing the decision rule upon the specification of error rates as Fisher and Neyman and Pearson had done, Wald gave formal recognition to the assessment of losses involved in making wrong decisions (Bahadur and Robbins, 1950; Lindley, 1953, 1961).

The formal mathematical structure of statistical decision theory is based upon the theory of games (Luce and Raiffa, 1957). It requires the statistical consumer or investigator to be able to list a set of possible actions, among which is a preferred action. The action preferred depends on what the "true state of nature" is, in other words, upon the unknown value of some parameter. The preferences among the actions are assessed in terms of the losses attached to each possible action to be taken in the presence of each state of nature (the circumstance which may be referred to as the loss function). The decision maker can obtain information about the state of nature at some specified cost for the observations made; he must balance cost against possible losses in the face of insufficient information. The decision function, then, depends upon the given loss function, the set of possible actions, the possible states of nature, the prior information available about the state of nature, and the cost of making observations (Savage, 1954).

Point estimation, interval estimation, and hypothesis testing are subsumed within a single theory, known as statistical hypothesis testing. The problem of choosing a design and an analysis is integrated within the theory and is closely associated with the factors determining the choice of the statistical decision function.

Another noteworthy contribution of Wald was his demonstration that there is not a unique decision procedure which is equally effective in all conditions. He showed that in those cases where prior information is available, a Bayes solution is admissible and optimum. There are situations, he pointed out, in which a Bayes solution is not part of an admissible strategy. He also raised the question of formulating a class of admissible hypotheses, and he thereby clarified some of the issues connected with composite hypotheses in the Neyman-Pearson formulation.

The theory of testing statistical hypotheses provides a decision procedure for working with multiple hypotheses rather than merely with two hypotheses, as in the theory of significance tests. For example, given three hypotheses, including a null hypothesis and two alternative hypotheses at different distances on either side of the null hypothesis, an optimum statistical decision function often is available to select the most probable of the three hypotheses on the available data (Lehmann, 1959).

The mathematical complexity of statistical decision theory and the great amount of information which is required to use it have discouraged its application in behavioral science research. Its theoretical formulations have helped to clarify a number of problems in the conventional use of significance tests (Kaiser, 1960).

Contributions of Bayesian Statistical Theory to Statistical Inference

Bayesian statistical theory, for the most part, is not so dependent on significance tests for informative statistical inferences as is classical statistical theory, namely the theories of Fisher and Neyman and Pearson. With rare exceptions, the British-American school of statistics has been firmly entrenched in the objective or frequency interpretation of probability and has eschewed the use of Bayes's theorem or the likelihood function as bases for formulating statistical inferences. The Continental school, on the other hand, has been concerned with giving formal mathematical treatment to the subjective theory of probability or with developing statistical methods based on the use of Bayes's theorem (Savage, 1954).

One British scientist working somewhat independently of the British-American school, his work unrecognized until recently by the majority of American statisticians, has made many important technical advances in the use of Bayes's theorem. Jeffreys (1957, 1961) has treated the problem of significance tests in a unique manner. Tests of significance are used to assess whether the hypothesized parameter(s) can be considered to account adequately for the obtained data. If the approximation of the posterior probability distribution to the null hypothesis, which reflects the prior probability distribution, is not satisfactory, then another parameter is introduced into the model. The process of adjusting the value of the null hypothesis is continued until a satisfactory approximation is attained, that is, within the limits of random sampling variation and errors of measurement. Jeffreys has emphasized the importance of refining measurement methods for the advancement of scientific knowledge. For Jeffreys one of the objectives of scientific investigation has been the verification of satisfactory null hypotheses. Perusal of the computations required in the use of Bayes's theorem for *t*-tests of significance would dismay the average social scientist accustomed to *t*-tests without consideration of prior probability distributions.

The merits of Jeffreys' work have been recognized by Lindley (1953, 1961), Raiffa and Schlaifer (1961), Savage (1954, 1961), and Savage and others (1962). His work and that of the French and Italian probability theorists have influenced the surgence of Bayesian statistics, especially Bayesian statistical decision theory in the United States (Raiffa and Schlaifer, 1961; Roberts, 1962; Schlaifer, 1959).

In his compendious discussion concerning the foundations of statistics of 10 years ago, Savage (Savage, 1954; Edwards, 1956) argued that various aspects of the classical theories of statistics could be reinforced and given a consistent logical framework relevant to the behavior of the scientist and the decision maker by adopting the personal interpretation of probability. At that time, he critically evaluated many of the conventional practices in the use of statistical methods, especially the widespread use of extreme null hypotheses which are known to be false prior to experimentation. Subsequently, Lindley (1961) and Raiffa and Schlaifer (1961) have built upon the work of Jeffreys a still incomplete structure of statistical theory and methods using Bayes's theorem. Edwards, Lindman, and Savage (1963) have presented to psychologists some of the fundamental ideas and techniques of Bayesian statistics.

Bayesian statisticians have pointed out that classical statistical methods do not make full use of available information and that they do not provide concise and relevant answers to the questions of greatest importance to investigators. Furthermore, in many situations the decision procedures do not meet the basic criterion of consistency (Pratt, 1961b).

Classical and Bayesian formulations have points in common. Both have the following elements: alternative hypotheses or acts, possible parameter values or states, the possibility of sampling (experimentation) to obtain information about the parameter value or state, the sampling or experimental results in the form of descriptive statistics, and either loss structures or error rates. Both seek the best decision rule which will minimize loss or error in the decisions or conclusions made following experimentation.

The classical decision maker chooses an act on the basis of the outcome of a sample which is conditional upon the parameter and the type of sample and experiment. Within the decision rules for significance tests, the so-called conditionality of the experiment and of the sample is not adequately represented; the statistician must use his judgment to give appropriate weight to the conditions when he interprets the obtained level of significance (Cox, 1958b).

The Bayesian statistician pursues a different course. He begins by making probabilistic statements about the parameter under investigation. The probabilistic statements are in the form of a prior distribution. Then, dependent upon how precise he wishes to be in his final probabilistic statements about the value of the parameter, he sets a limit, or an upper bound, in the form of an expected value on the cost or worth of experimental data for selecting a terminal act or decision. From this point of view he decides whether it is worthwhile to sample. As sampling information becomes available, the Bayesian modifies his prior distribution in the light of sample evidence and thereby obtains a posterior distribution. By using some specified decision rule, he may decide that he has sufficient information to come to a decision or a conclusion; or he may decide that his uncertainty has not been reduced enough and that, therefore, he must continue sampling until he has a suitable basis for a terminal act or conclusion (Raiffa and Schlaifer, 1961; Roberts, 1962).

The Bayesian analysis formalizes many aspects of statistical practice which classical methods leave to the judgment of the investigator. For the orthodox contention that there is often little objective basis for arriving at prior probability distributions, Bayesians have countered that in at least a number of situations the prior distribution assumes little weight in the final or posterior distribution (Edwards, Lindman, and Savage, 1963; Savage and others, 1962). Several interesting solutions to this problem have been indicated by Raiffa and Schlaifer (1961).

Edwards, Lindman, and Savage (1963), Lindley (1961), and Savage and others (1962) have discussed some of the advantages of Bayesian significance tests over classical tests. The first-mentioned authors have discussed the importance in many educational and psychological research problems of taking possible losses or gains and prior information into account. For example, classical *t*-tests and *F*-tests have been used to study the differential effects on several groups of several methods of instruction. In neither approach has it been practical to take into account considerations of previously available information about the differential effectiveness of one method versus another; nor has it been customary to give attention to various kinds of losses, economic or learning, which might be associated with the adoption of one method rather than another. The authors have showed how Bayesian procedures can provide appropriate answers to such problems.

The Likelihood Principle and Statistical Inference

The likelihood principle, like Bayes's theorem, has been recognized as a part of the armamentarium of statistics; but, as with Bayes's theorem, it has been given little attention as a primary tool for statistical inference. The likelihood function is represented by the distribution function of the random sample variables corresponding to values of given parameters and is determined by the observed outcome of a random variable in any specified experiment (Birnbaum, 1962). Barnard, Fisher, and Birnbaum (Birnbaum, 1962) have argued that the likelihood function alone provides a suitable and sufficient basis for interpreting experimental data. The likelihood function can be interpreted without reference to the structure of the experiment. In contrast with the classical methods of significance tests and confidence levels, its use avoids the problem of having to modify the interpretation of the significance and confidence levels in the light of the experimental frame of reference. The likelihood function clearly asserts the irrelevance of possible experimental outcomes which have not been observed during any given experiment. One of the main objections to the use of the tail areas of the probability distribution for significance tests is just that unobserved experimental possibilities are irrelevant for the formulation of inferences about the actual observations. The likelihood function and Bayes's theorem are not dependent upon the sequence of or the stopping point in sampling, whereas orthodox significance tests and confidence intervals cannot be used when sampling is stopped arbitrarily.

Use of Significance Tests for Statistical Inference

The test of significant difference has been called the prototype of modern experimental statistics (Tukey, 1960a). It is a qualitative rather than a quantitative procedure for statistical analysis and inference which is used to answer questions such as this: Dare we conclude that the difference is not zero? A classical significance test assays whether the two statistical hypotheses A and B are equal or whether A is less than or greater than B. The failure to obtain a significant difference does not warrant the conclusion that A is equal to B until careful consideration is given to the specific experimental situation, the available evidence, and the assessment of the consequences for such a decision or conclusion (Tukey, 1960a). The precision of the experimental comparison, the power of the statistical test, and the theoretical and the empirical closeness of A to B must be examined before the lack of statistical significance is interpreted as a positive finding (Cox, 1958a; Tukey, 1960a). The effectiveness of confidence intervals in showing the probable relationship between A and B indicates their superiority over significance tests for informative inference. Statisticians have given a great deal of attention to the use and the misuse of significance tests (Anscombe, 1961; Bahadur and Robbins, 1950; Good, 1958; Kish, 1959; Lindley, 1958; Savage, 1957; Selvin, 1957; Sterling, 1959, 1960; Williams, 1959). There has been general agreement that significance tests have been used too frequently in social science research (Grant, 1962; Harrington, 1961; Lubin, 1962; Nunnally, 1960; Rozeboom, 1960; Savage, 1957; and Wilson, 1961) and at the expense of more appropriate procedures, such as confidence and fiducial intervals (Birnbaum, 1961; Grant, 1962; Kish, 1959; Lindley, 1958; and Savage and others, 1962). Significance tests without power functions do not answer such questions as these: How far is the sample statistic from the null hypothesis? How much credence should be given to the null or the alternative hypotheses?

The situations in which classical significance tests can be used are limited. Anscombe (1961) has pointed out the usefulness of significance tests as a fundamental method for testing theoretical hypotheses or models. For this purpose, the null hypothesis is given some credence in the light of theoretical and experimental consideration. The investigator wishes to determine in a new sample whether the null hypothesis is tenable, the observed departures being no more than those expected as a result of random sampling variations. When the departures are excessive, the investigator considers whether another parameter should be introduced into the statistical model or whether some other modification in the model should be made. Another use for significance tests occurs when an experimenter wishes to test the adequacy of a particular statistical design, an analysis or technique, a particular stochastic process, or an experimental procedure. He is interested in determining whether the methods used produce the desired precision or uniformity with due allowance for random sampling errors (Anscombe, 1961). Significance tests are also applicable when an investigator is concerned with testing a particular statistical hypothesis, when a set of specific alternatives has not yet been conceptualized (Savage and others, 1962). When a research worker wishes to verify a prediction that an experimental result has been in a specific direction, one-tailed significance tests are appropriate (Anscombe, 1961).

Significance tests can be informative in later stages of research when estimation is included in the statistical procedures. The null hypothesis of no difference has been judged to be no longer a sound or fruitful basis for statistical investigation. Both the null and the alternative hypotheses are formulated to include estimation of relevant experimental variables (Bush, 1963; Grant, 1962; Nunnally, 1960; Tukey, 1960b).

Significance tests can be divided into two groups: those for which a set of alternative hypotheses can be defined, such as tests of specific parameter values, and those for which a set of mutually exclusive and exhaustive hypotheses are not available, such as tests of randomness. The former almost always have comparable interval estimation procedures—at least, such procedures can be derived—whereas the latter pose serious problems for the derivation of interval estimation procedures. When comparable interval estimation procedures are not available, then an investigator has no choice but to use significance tests, as in the use of some distribution-free methods (Kendall and Stuart, 1961; Savage and others, 1962).

There are many statistical problems for which both significance and interval estimation procedures are available. Similar answers are obtained if the procedures are used effectively and appropriately (Berkson, 1942; Birnbaum, 1961; Kendall and Stuart, 1961; Natrella, 1960; Pratt, 1961a). Statisticians have agreed that interval estimation does give a more intuitively understandable summary of the data than do significance tests with power functions (Kish, 1959; Natrella, 1960).

Among statistical problems which often can be treated by either testing or interval estimation procedures are (a) simple preference problems, wherein a family of simple hypotheses is ranked in order of credibility on the basis of the data, and (b) composite preference problems, wherein a family of composite hypotheses is ranked on the basis of the data (Lehmann, 1959; Savage and others, 1962). These problems conventionally have been classified under the rubrics of hypothesis testing, point and interval estimation, and discrimination (Savage and others, 1962).

Problems in Statistical Hypothesis Testing

There are many unsolved problems in the theory and the practice of statistical hypothesis testing. Some of the theoretical problems have been mentioned at the beginning of this chapter. In the following paragraphs some of the practical problems are summarized.

Decision Theory and Conclusion Theory

Tukey (1960a) has suggested that some of the misunderstanding surrounding the use of significance tests can be clarified by distinguishing those statistical problems concerned with the necessity of making a choice between two alternatives from those concerned with the accumulation of evidence in support of a theory or set of working hypotheses. Decision theory is typically concerned with problems in which economic losses or rewards, a set of possible actions, and their consequences in various states of nature can be defined. Decision theory counsels the statistical consumer about how to choose wisely among available strategies (Bahadur and Robbins, 1950; Flanagan, 1958).

Conclusion theory, on the other hand, is concerned with evaluating the adequacy of evidence. Conclusions need not be made if the evidence is deemed inadequate. A conclusion is accepted relative to the conditions of an experiment until compelling evidence to the contrary is found. Decision procedures choose between two hypotheses in terms of minimizing the risk for an action, while conclusion procedures often are concerned with controlling errors of both the first and the second kinds at suitably low levels in choosing between two hypotheses (Anscombe, 1961; Lindley, 1961; Tukey, 1960a).

Rejection or Acceptance of the Null Hypothesis

Much of the controversy among statisticians and behavioral scientists about the theory and the practice of statistical hypothesis testing has centered upon whether the null hypothesis can reasonably be accepted. Grant (1962) has argued that experiments oriented toward the acceptance of null hypotheses are inappropriate. Binder (1963) has countered that the Neyman-Pearson formulation admits the validity of accepting the null hypothesis, although, unless some adjustment in significance level is made, the rationale is prone to reject the null hypothesis when sample sizes are large or when powerful tests are used (Pratt, 1961b). It is the ready rejection of the null hypothesis by classical methods that is the focus of much of the Bayesian criticism (Edwards, Lindman, and Savage, 1963; Savage and others, 1962).

Among statisticians, Anscombe (1961), Good (1958), Lindley (1958, 1961), Savage and others (1962), and Sterling (1960) have recommended that hypothesis testing should be directed toward testing plausible null hypotheses and not toward testing implausible ones that can be rejected by a single observation.

Statistical Significance and Substantive Significance

That "significance levels do not signify" (Savage and others, 1962) has been widely recognized as a primary problem in the use of significance tests. Among the proposed solutions to the problem have been the inclusion of a distance function to indicate how deviant the data are from the null hypothesis (Bulmer, 1957), the use of a form of variance in analysis of variance to show how well the independent variables account for the obtained results (Bolles and Messick, 1958; Gaito, 1958), the substitution of interval estimation for significance tests to give a quantitative representation of the spread of the obtained values (Birnbaum, 1961; Kish, 1959; Natrella, 1960), the use of Bayesian methods (Edwards, Lindman, and Savage, 1963), and the use of the simple likelihood function (Birnbaum, 1962).

Significance tests as methods for informative inference have been criticized severely (Birnbaum, 1962; Edwards, Lindman, and Savage, 1963; Lubin, 1962; Rozeboom, 1960; Savage and others, 1962). Mere inspection of the data often reveals the untenability of the null hypothesis. When there are grounds for believing that random sampling variations alone do not account for the data, a significance test is not the most appropriate method to use (Williams, 1959). Significance tests do not summarize the evidence (Birnbaum, 1962; Lindley, 1958; Tukey, 1962; Wilson, 1961). Much caution must be used in interpreting significance levels, for the obtained level of significance is not consistent from sample to sample or among comparable tests (Pratt, 1961b). The obtained level of significance is dependent upon conditions which do not modify the obtained level (Birnbaum, 1962). There are many conditions which vitiate the accuracy of the obtained significance levels, even for robust test procedures (Tukey, 1962). Often graphical representation indicates the presence of disturbing conditions, but social scientists do not use such a procedure regularly (Tukey, 1962). Classical significance procedures are not sensitive to the size of the error (Wilson, 1961); often it is the size of the error that an investigator wishes to know to assess his procedures.

An artificial level of significance is not a sufficient or an efficient indication of the presence or absence of evidence; yet editorial policies have chosen to ignore this fact in their insistence on extreme levels of significance (Melton, 1962). Publication policies have posed other problems for the interpretation of test results. Sterling (1959) has commented upon the prevalence of Type I errors among psychological publications associated with the suppression of nonsignificant results as a consequence of the inclination of editors to publish mainly articles with statistically significant findings. Cohen (1962) has found that articles on social and abnormal psychology have not been appropriately concerned with the power of the designs and tests used.

Significance Tests and Statistical Assumptions

The assumptions underlying all phases of statistical methods constitute pervasive problems in statistical practice (Savage, 1957; Savage and others, 1962). Rarely can an investigator be certain that the assumptions are adequately fulfilled in any particular problem (Tukey, 1960a, 1962). Among the important problems regarding assumptions in ordinary statistical practice are the following: (a) the appropriateness of the probability model to characterize the population (Savage, 1957; Neyman, 1960): (b) the inappropriate use of small sample methods before large sample methods have been used to make estimates of variance (Nunnally, 1960); (c) the robustness or sensitivity of statistical methods to violations of such assumptions as normality, independence, homogeneity, additivity (Boneau, 1960); (d) the selection of test procedures with due awareness for the performance and the requirements of the procedures and the nature of the data (Binder, 1959; Boneau, 1960; Kendall and Stuart, 1961; Lehmann, 1959); (e) the use of transformations of data to meet the assumptions of the more powerful statistical methods versus the use of the less powerful distribution-free and nonparametric methods (Boneau, 1962; Lehmann, 1959; Savage, 1957). Always the investigator must remember that different tests based on different assumptions when applied to varying sets of data lead to different results (Tukey, 1962). Often the precision promised is not the precision delivered (Tukey, 1954, 1962).

Another basic assumption underlying classical statistical methods is that of randomization. Psychologists have been criticized for not being sufficiently concerned with random sampling procedures, the definition of target populations, and the limitations in the representative nature of most samples studied (Kish, 1959; McGinnis, 1958; Nunnally, 1960; Stevens, 1960). McGinnis (1958) and Kish (1959) have discussed the validity of significance tests in surveys versus well-controlled experiments.

Data Analysis and Statistical Hypothesis Testing

Tukey (1962) commented that much of current statistical practice has emphasized the development of elaborate methods at the expense of attention to the analysis of data. Instead of concentrating upon probing the nature of the data and upon the questions that concern scientists, statistical practitioners have been prone to commit errors of the third kind—that is, giving exact answers to the wrong questions—which is perhaps the most serious of the three kinds of errors (Schlaifer, 1959). The routine use of significance tests has perpetrated many errors of the third kind, for, as mentioned above, significance tests do not provide the information that scientists need, and, furthermore, they are not the most effective method for analyzing and summarizing data. Tukey (1954, 1962) recommended that attention should be given to adapting methods to problems rather than to forcing problems into particular methods.

Statistics and Experimentation

Statistical methods have been widely used in the sciences in which strict experimental controls are either not feasible or desirable. The usefulness of statistical methods for designing and analyzing investigations has been widely recognized; however, there have been dissenters, among them Hogben (Hogben, 1957; Stevens, 1960), who have maintained that the most appropriate solution for the many problems associated with the theory and practice of statistics is the replacement of statistical methods by well-controlled experimentation. Lindley (1958), in his response to Hogben's criticisms, contended that statistical methodology requires extensive modifications in order to play an effective part in science, but that statistics will doubtlessly continue to have an important role in all kinds of research.

The problems of the relative emphasis which should be given to considerations of statistical versus experimental design does not admit an easy solution.

Statistics and Psychological Theory

If statistics is to make significant contributions to the testing of scientific hypotheses, then the methodology of statistical hypothesis testing must incorporate theoretical formulations of the subject matter under investigation. Statistical inferences are maximally informative if they are supported by other findings and if they are enmeshed in an expanding network of hypotheses and constructs.

The constitution of a set of mutually exclusive and exhaustive statistical hypotheses requires careful consideration of relevant theoretical alternatives in order to select heuristically valuable alternatives (Savage and others, 1962). Recognition of the limitations in the theoretical formulations and in the statistical methods should motivate prudence against unwarranted statistical and scientific inferences (Jeffreys, 1957).

Bush (1963) remarked that theories are not substantiated by the goodness of fit between the data and a hypothetical statistical model but rather by the estimation of significant parameters in the statistical model so that information about the functions of experimental and theoretical variables can be accumulated. Estimation is essential for the testing of accurate and fruitful theories.

Conclusion

While the so-called crisis in statistics (Hogben, 1957; Lindley, 1958; Stevens, 1960) has not been resolved, it has highlighted many of the shortcomings in such frequently used methods as significance tests, and it has evoked constructive thought about the very foundations of statistical inference (Savage and others, 1962) as well as about the rationale of statistical methods (Tukey, 1954, 1962).

If educational and psychological research workers heed the admonitions and the recommendations of statisticians, a change in many aspects of statistical practices may be anticipated. At least investigators will be wary of the routine application of significance tests as the main basis for statistical inference.

Bibliography

Anscombe, F. J. "Examination of Residuals." *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press, 1961. Vol. 1, pp. 1-35.

Bahadur, Raghu Raj, and Robbins, Herbert. "The Problem of the Greater Mean." *Annals of Mathematical Statistics* 21: 469-87; December 1950.

Berkson, Joseph. "Tests of Significance Considered as Evidence." *Journal of the American Statistical Association* 37: 325-35; September 1942.

Binder, Arnold. "Considerations of the Place of Assumptions in Correlational Analysis." *American Psychologist* 14: 504-10; August 1959.

Binder, Arnold. "Further Considerations on Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 70: 107-15; January 1963.

Birnbaum, Allan. "Confidence Curves: An Omnibus Technique for Estimation and Testing Statistical Hypotheses." *Journal of the American Statistical Association* 56: 246-49; June 1961.

Birnbaum, Allan. "On the Foundations of Statistical Inference." *Journal of the American Statistical Association* 57: 269-326; June 1962.

Bolles, Robert C. "The Difference Between Statistical Hypotheses and Scientific Hypotheses." *Psychological Reports* 11: 639-45; December 1962.

Bolles, Robert, and Messick, Samuel. "Statistical Utility in Experimental Inference." *Psychological Reports* 4: 223-27; June 1958.

Boneau, C. Alan. "The Effects of Violations of Assumptions Underlying the t Test." *Psychological Bulletin* 57: 49-64; January 1960.

Boneau, C. Alan. "A Comparison of the Power of the U and t Tests." *Psychological Review* 69: 246-56; May 1962.

Buehler, Robert J. "Some Validity Criteria for Statistical Inferences." *Annals of Mathematical Statistics* 30: 845-63; December 1959.

Bulmer, M. G. "Confirming Statistical Hypotheses." *Journal of the Royal Statistical Society* (Series B, Methodological) 19: 125-32; No. 1, 1957.

Bush, Robert R. "Estimation and Evaluation." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, Chapter 8, pp. 429-69.

Cohen, Jacob. "The Statistical Power of Abnormal-Social Psychological Research: A Review." *Journal of Abnormal and Social Psychology* 65: 145-53; September 1962.

Cox, D. R. *Planning of Experiments.* New York: John Wiley & Sons, 1958. 308 pp. (a)

Cox, D. R. "Some Problems Connected with Statistical Inference." *Annals of Mathematical Statistics* 29: 357-72; June 1958. (b)

Edwards, Ward. "Savage Statistics." *Contemporary Psychology* 1: 14-15; January 1956.

Edwards, Ward; Lindman, Harold; and Savage, Leonard J. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70: 193-242; May 1963.

Fisher, Sir Ronald A. *Statistical Methods and Scientific Inference.* New York: Hafner Publishing Co., 1956. 175 pp.

Fisher, Sir Ronald A. *The Design of Experiments.* Seventh edition. New York: Hafner Publishing Co., 1960. 248 pp.

Flanagan, John C. "The Dollar Dimension in Testing Decisions." *Contemporary Psychology* 3: 164-66; June 1958.

Gaito, John. "The Bolles-Messick Coefficient of Utility." *Psychological Reports* 4: 595-98; December 1958.

Good, I. J. "Significance Tests in Parallel and in Series." *Journal of the American Statistical Association* 53: 799-813; December 1958.

Grant, David A. "Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 69: 54-61; January 1962.

HARRINGTON, GORDON M. "Statistics' Logic." *Contemporary Psychology* 6: 304-305; September 1961.

HOGBEN, LANCELOT. *Statistical Theory: The Relationship of Probability, Credibility, and Error.* New York: W. W. Norton & Co., 1957. 510 pp.

HOTELLING, HAROLD. "The Impact of R. A. Fisher on Statistics." *Journal of the American Statistical Association* 46: 35-46; March 1951.

HOTELLING, HAROLD. "The Statistical Method and the Philosophy of Science." *American Statistician* 12: 9-14; December 1958.

JEFFREYS, HAROLD. *Scientific Inference.* Second edition. New York: Cambridge University Press, 1957. 236 pp.

JEFFREYS, HAROLD. *Theory of Probability.* Third edition. New York: Oxford University Press, 1961. 447 pp.

KAISER, HENRY F. "Directional Statistical Decisions." *Psychological Review* 67: 160-67; May 1960.

KENDALL, MAURICE G., and STUART, ALAN. *The Advanced Theory of Statistics.* Revised edition. New York: Hafner Publishing Co., 1961. Vol. 2, "Inference and Relationship," 676 pp.

KISH, LESLIE. "Some Statistical Problems in Research Design." *American Sociological Review* 24: 328-38; June 1959.

LEHMANN, E. L. *Testing Statistical Hypotheses.* New York: John Wiley & Sons, 1959. 369 pp.

LINDLEY, D. V. "Statistical Inference." *Journal of the Royal Statistical Society* (Series B, Methodological) 15: 30-76; No. 1, 1953.

LINDLEY, D. V. "Professor Hogben's 'Crisis'—A Survey of the Foundations of Statistics." *Applied Statistics* 7: 186-98; November 1958.

LINDLEY, D. V. "The Use of Prior Probability Distributions in Statistical Inference and Decisions." *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press, 1961. Vol. 2, pp. 453-68.

LUBIN, ARDIE. "Statistics." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1962. Vol. 13, pp. 345-70.

LUCE, R. DUNCAN, and RAIFFA, HOWARD. *Games and Decisions: Introduction and Critical Survey.* New York: John Wiley & Sons, 1957. 509 pp.

MCGINNIS, ROBERT. "Randomization and Inference in Sociological Research." *American Sociological Review* 23: 408-14; August 1958.

MCNEMAR, QUINN. "At Random: Sense and Nonsense." *American Psychologist* 15: 295-300; May 1960.

MELTON, ARTHUR W. "Editorial." *Journal of Experimental Psychology* 64: 553-57; December 1962.

MOWRER, O. HOBART. *Learning Theory and the Symbolic Processes.* New York: John Wiley & Sons, 1960. 473 pp.

NATRELLA, MARY G. "The Relation Between Confidence Intervals and Tests of Significance." *American Statistician* 14: 20-22, 38; February 1960.

NEYMAN, JERZY. "Indeterminism in Science and New Demands on Statisticians." *Journal of the American Statistical Association* 55: 625-39; December 1960.

NUNNALLY, JUM. "The Place of Statistics in Psychology." *Educational and Psychological Measurement* 20: 641-50; Winter 1960.

PEARSON, E. S. "Some Thoughts on Statistical Inference." *Annals of Mathematical Statistics* 33: 394-403; June 1962.

PRATT, JOHN W. "Length of Confidence Intervals." *Journal of the American Statistical Association* 56: 549-67; September 1961. (a)

PRATT, JOHN W., reviewer. "*Testing Statistical Hypotheses* by E. L. Lehmann." *Journal of the American Statistical Association* 56: 163-67; March 1961. (b)

RAIFFA, HOWARD, and SCHLAIFER, ROBERT. *Applied Statistical Decision Theory.* Boston: Division of Research, Harvard Business School, Harvard University, 1961. 356 pp.

ROBERTS, HARRY V., reviewer. "*Applied Statistical Decision Theory* by Howard Raiffa and Robert Schlaifer." *Journal of the American Statistical Association* 57: 199-202; March 1962.

ROZEBOOM, WILLIAM W. "The Fallacy of the Null-Hypothesis Significance Test." *Psychological Bulletin* 57: 416-28; September 1960.

Savage, I. Richard. "Nonparametric Statistics." *Journal of the American Statistical Association* 52: 331-44; September 1957.

Savage, Leonard J. *The Foundations of Statistics.* New York: John Wiley & Sons, 1954. 294 pp.

Savage, Leonard J. "The Foundations of Statistics Reconsidered." *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press, 1961. Vol. 1, pp. 575-86.

Savage, Leonard J., and others. *The Foundations of Statistical Inference.* New York: John Wiley & Sons, 1962. 112 pp.

Schlaifer, Robert. *Probability and Statistics for Business Decisions.* New York: McGraw-Hill Book Co., 1959. 732 pp.

Selvin, Hanan C. "A Critique of Tests of Significance in Survey Research." *American Sociological Review* 22: 519-27; October 1957.

Sterling, Theodor D. "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance—or Vice Versa." *Journal of the American Statistical Association* 54: 30-34; March 1959.

Sterling, Theodor D. "What Is So Peculiar About Accepting the Null Hypothesis?" *Psychological Reports* 7: 363-64; October 1960.

Stevens, S. S. "The Predicament in Design and Significance." *Contemporary Psychology* 5: 273-76; September 1960.

Tukey, John W. "Unsolved Problems of Experimental Statistics." *Journal of the American Statistical Association* 49: 706-31; December 1954.

Tukey, John W. "Conclusions vs. Decisions." *Technometrics* 2: 423-33; November 1960. (a)

Tukey, John W. "Where Do We Go from Here?" *Journal of the American Statistical Association* 55: 80-93; March 1960. (b)

Tukey, John W. "The Future of Data Analysis." *Annals of Mathematical Statistics* 33: 1-67; March 1962.

Wald, Abraham. *Statistical Decision Functions.* New York: John Wiley & Sons, 1950. 179 pp.

Williams, E. J. *Regression Analysis.* New York: John Wiley & Sons, 1959. 214 pp.

Wilson, Kellogg V. "Subjectivist Statistics for the Current Crisis." *Contemporary Psychology* 6: 229-31; July 1961.

Yates, F. "The Influence of *Statistical Methods for Research Workers* on the Development of the Science of Statistics." *Journal of the American Statistical Association* 46: 19-34; March 1951.

CHAPTER II

Selected Contributions to Parametric and Nonparametric Statistics

WILLIAM B. MICHAEL

ALTHOUGH THERE is a slight duplication of the research reviewed in this chapter with that to be found in Chapters I and III, an effort has been made to include those contributions to statistical methodology that were thought to be of at least some interest and of some potential help to research workers in the behavioral sciences. The scope of this particular chapter is probably broader and perhaps less unified than that of any other in this issue of the REVIEW. In this chapter, the major objective is to furnish for the period from July 1, 1960, to July 1, 1963, a brief description both of the general developments in statistical inference and of the somewhat specific contributions that have been made to the theory and application of parametric and nonparametric tests of significance—especially those tests involving only two samples. Consideration has also been given to citing important books, to describing advances realized in the treatment of data in contingency tables, and to mentioning certain developments in the methodology of multiple comparisons that have not been noted by Norton in Chapter III.

Books

As one might expect, scores of books on statistical methodology have appeared since July 1960. In this chapter, those books will be mentioned which for the most part are somewhat general in their orientation. For example, books will not be cited which are concerned primarily with such specific problem areas as analysis of variance and covariance, regression and correlation, factor analysis, sample surveys, scaling and measurement theory, statistical decision theory and utility measurement, statistical learning theory, and computer methodology including numerical analysis, as the authors of the other chapters have made reference to most of the important books in these specific areas.

For research workers in education and in the behavioral sciences, the third revision of the well-known text by McNemar (1962), though not much changed, remains as a standard reference text. For a similar audience, the relatively elementary text by Peatman (1963) should also be useful. Of particular interest to the behavioral scientist is Horst's (1963) lucid text concerning matrix algebra. For the educational research worker who wishes an introduction to probability and statistical applications in terms of a modern point of view embodying elementary set theory, the volume by Mosteller, Rourke, and Thomas (1961) should be especially helpful.

Requiring a relatively elementary level of mathematical sophistication, about the equivalent of one or two years of college mathematics, the texts by Alexander (1961), Freund (1960, 1962), and Kozelka (1961) are noteworthy. Demanding considerably less mathematical maturity of the reader than the four books just cited is a text by Hammond and Householder (1962), which rests heavily upon a highly verbalized, nonmathematical approach to the understanding of statistics.

At a high level of mathematical sophistication are the following works: the second volume of Kendall and Stuart's (1961) *The Advanced Theory of Statistics*, which is concerned with statistical problems of inference and relationship; Walsh's (1962) *Handbook of Nonparametric Statistics;* and Wilks's (1962) *Mathematical Statistics.* Probably the most profound treatise concerning the nature of probability is the revised edition of the volume by Jeffreys (1961), who assumed the point of view of a philosopher of science. Constituting a monumental contribution to the conceptual framework of probability and statistics is the short but penetrating monograph—a collection of papers—by Savage and others (1962) titled *The Foundations of Statistical Inference*, to which reference will be made in the next section.

Additional References: Sarhan and Greenberg (1962); Savage (1962).

General Contributions to the Theory of Statistical Inference

Three Basic Contributions

In Chapters I and III, Clark and Norton have given attention both to Birnbaum's (1962) general theoretical paper on the foundations of statistical inference, in which the likelihood function is the unifying principle, and to Tukey's (1962) comprehensive paper on the future of data analysis. The accompanying discussion and comments of 14 renowned statisticians concerning Birnbaum's formulation are particularly informative. In Chapter I, Clark has cited 20 times the important monograph by Savage and others (1962), in which Savage not only argued in considerable detail for a theory of statistical inference based upon Bayes's theorem but also answered questions from a number of illustrious British statisticians who had presented alternative points of view in their short papers. Anyone in educational and psychological research who is seriously interested both in trying to understand the conceptual foundations of statistical inference and in attempting to employ wisely the tests of statistical methodology should read and reread with great care these three significant contributions. Admittedly, there will be much that the majority of quantitatively oriented research workers in both education and psychology will not grasp to their complete satisfaction in view of the uneven nature of the demands placed upon the level of mathematical maturity of the reader. Nevertheless, there are many key concepts, basic points of view, and evaluative comments that are not only quite comprehensible but also helpful in broadening one's orientation to the basic problems of statistical methodology.

Contributions from AERA-Sponsored Publications

Within the field of education itself, two general presentations concerning statistical methodology with particular emphasis upon problems of inference may be found in AERA-sponsored publications. In the *Encyclopedia of Educational Research,* Harris (1960) furnished a lucid and succinct presentation of key problem areas in statistical inference. From the standpoint of the place which statistics plays in the application of the scientific method to research on teaching, the chapter by Tatsuoka and Tiedeman (1963) in the *Handbook of Research on Teaching,* edited by Gage (1963), should be most helpful to its intended audience. The authors considered a number of statistical techniques and furnished a helpful table that with its many panels serves to organize these statistical techniques according to role, to the type of scale required for appropriate application, and to the number of independent and dependent variables present.

Contributions from Psychological Literature

Since July 1960, at least six important expository articles have been concerned primarily with problems of statistical inference in psychology. In the main, they have been written in such a manner that most research-oriented psychologists and educational psychologists, irrespective of area of major interests, may be expected to understand a substantial portion of the material presented. Since, in Chapter I, Clark has referred to each of these contributions several times and has evaluated major portions of their contents, each one will be cited only briefly—primarily for the purpose of furnishing a relatively complete set of references in the general area of statistical inference. For the period from June 1959 to April 1961, Lubin (1962) reviewed more than 130 contributions from the statistical literature and emphasized inferential developments. His presentation, which is simultaneously critical and humorous, deserves careful study.

In his straightforward presentation of the place of statistics in psychology, Nunnally (1960) questioned the emphasis that has been given to the null hypothesis. Similarly, in what may be viewed as a possible landmark for developing a much-needed skepticism on the part of psychologists toward the use of the null hypothesis, Rozeboom (1960) vigorously criticized what he termed the "null-hypothesis decision procedure" as being an inappropriate method of inference. Relative to the development of a strategy of experimental investigation of theoretical models, Grant (1962) expressed doubts regarding the value of testing the null hypothesis when theoretical considerations lead to a prediction of differences of specified amount, and he then recommended a procedure involving estimation of confidence intervals with respect to the empirical means (averages) that can be referred to a theoretical curve. Taking a somewhat contrary position to that of Grant, Binder (1963) emphasized that the optimal design of experiments with sufficient built-in sensitivity to allow for acceptance of useful theories and for rejection of unsuitable ones is the key point rather than Grant's argument concerning the inappropriateness of an experimental design in which acceptance of the null hypothesis will suggest support for a theory.

Of marked importance to educational and psychological research workers was the comprehensive and definitive paper by Edwards, Lindman, and Savage (1963) concerning the use of Bayesian statistical inference in psychological research. They pointed to the fact that evidence obtained by the Bayesian statistician may frequently support the acceptance of the null hypothesis, whereas use of standard statistical procedures would lead to the rejection of the null hypothesis.

Contributions from Mathematical Statisticians

From the mathematical statisticians, three papers largely descriptive in nature deserve brief mention. In an expository article which reflected developments in inferential statistics over a period of nearly 30 years, Pearson (1962) evaluated the work which he and Neyman had done concerning confidence intervals, graciously acknowledged the contributions of other statisticians such as Fisher with differing points of view, and expressed certain reservations about subjectively considered losses and prior distributions of Bayesian statistics. Brillinger (1962) explored the general requirements for application of the fiducial argument, showed relationships to Bayesian theory, presented several illustrative examples, and included a comprehensive bibliography of 54 references. In a theoretical paper, Fraser (1961) reviewed several problems that have been analyzed in terms of the fiducial method, discussed a number of findings arising from these problems, and developed an abstract mathematical framework within which probability concepts could be related to a large class of problems.

Theory of Estimation

Point Estimation

Probably the most noteworthy paper concerning the theory of estimation was that by Birnbaum (1961), who extended and unified several previous formulations and points of view regarding estimation for one-parameter problems. His basic criterion was that of admissibility of a point estimator, which was considered in terms of its full distribution rather than in terms of its special loss functions. Both theoretical bases for characterizing admissible estimators and practical computational methods were included. Moreover, Birnbaum succeeded in generalizing Fisher's maximum likelihood principle of estimation, gave it exact (or nonasymptotic) justification, unified it with a theory of tests and the theory of Neyman-Pearson confidence regions. and described relationships between exact and asymptotic results.

To furnish an estimate of the common mean from two independent sample means as well as estimates of the separate variances of two normal populations, Fields, Kramer, and Clunies-Ross (1962) described an iterative method that yields either maximum likelihood values or quantities only slightly different from them. In terms of nonparametric rank test statistics, Hodges and Lehmann (1963) explored different approaches to point estimation, defined point estimates of location or shift in parameter, described their success in providing robust power. and proved certain properties of invariance and regularity.

Interval Estimation

During the past three years, much attention has been placed upon the length of confidence intervals. In the determination of the size of a sample in order that there will be a specified probability $1 - \alpha$ that a confidence interval contain a parameter and that there be a specified probability that the width of the confidence interval will be less than a certain designated amount, Graybill and Morrison (1960) outlined two requirements that must be met: (a) knowledge of the form of the frequency function from which a sample of size n is taken and (b) availability of previous information regarding the unknown parameter in the distribution. They suggested that a first sample be employed to determine the number of observations to be taken in a second sample in order that the two requirements may be satisfied.

Since, at most, only one point within a confidence interval can be the parameter, one possible measure of the desirability of a confidence interval procedure is the extent to which an interval contains false values as reflected by its expected length, since this length is an indication of the "average extent" to which false values are included. Embarking upon a new approach, Pratt (1961) attempted to determine the probability that a confidence interval procedure contains a *particular* false value through integrating over false values. In demonstrating that the expected length of a confidence interval is equal to the integral over false values of the probability that each false value is included, he noted that his approach yielded an equivalent measure for choosing among confidence interval procedures and presented numerical examples to illustrate his point. In a subsequent article based upon the work of Pratt, Madansky (1962) took a specific example, applied two different methods of obtaining a confidence interval for a parameter, found two different values for the interval, and showed that for one-sided intervals the expected length of a confidence interval and the probability of covering false values were not related.

In other papers, Halperin (1961) discussed confidence intervals in the instance of censored samples; Harter (1960b) furnished tables of percentage points for the range, moments of the range, and the studentized range; Blyth and Hutchinson (1960) presented a tabulation of the .95 and .99 confidence intervals of the binomial parameter that were based on Neyman's shortest unbiased estimates; and Pachares (1960) also developed extensive tables of confidence limits for the binomial distribution for values of .005, .010, .025, and .050 in a.

Additional References: Birnbaum (1961); Harter (1961, 1963); Murthy (1962); Pachares (1961).

Inference Concerning Authorship

In a paper of considerable potential interest to historians and linguistic analysts who may wish to infer the authorship of documents and papers, Mosteller and Wallace (1963) undertook an investigation in which they compared discrimination methods, carried out explorations with Bayesian approaches in large-scale data analysis, attempted to ascertain whether Madison or Hamilton wrote the disputed *Federalist* papers, and suggested routine ways for solving problems of authorship. They noted that the frequency of appearance of articles, prepositions, and conjunctions, which occur at fairly stable rates, yield important discriminative information. On the other hand, many meaningful nouns related to the subject matter at hand do not provide stable rates of frequency per page. Mosteller and Wallace reported odds that were highly in favor of Madison's having written all 12 of the *Federalist* papers that have been in dispute. Use of different sets of prior probabilities was considered to be a worthwhile procedure in the inferential process.

To those certain college and university administrative and faculty review committees that are responsible for making recommendations concerning whether professors should be promoted in rank and/or should be given increases in salary, the techniques proposed by Mosteller and Wallace may be of considerable interest—especially in the instance in which the reprints presented by a professor as evidence of his scholarly attainment list two or more authors. These techniques would afford a committee with a crusading or crucifying spirit a means for inferring whether the professor under consideration indeed did undertake most of the writing appearing in his joint publications or whether his collaborators actually prepared the text. Of course, samples of previously published writings by each coauthor would be necessary to furnish a base for comparison.

In the event that the reader, who as a coauthor himself may be struggling in the publish-and/or-perish reprint race, is beginning to tremble at the thought of possible exposure, this writer does not wish to frighten him unduly, for it is highly unlikely that any senior faculty committee member or any administrative officer would dare take precious time from, respectively, his own zestful scholarly activities or his challenging paperwork responsibilities for such an unrewarding inquisition. Then, too, what is the probability that either a college administrator or anyone with the statistical competence to do such an inferential study would be reading this chapter? It would seem safe to say that the probability is low—exceedingly low indeed ($p < .01$).

Significance Tests

Parametric Versus Nonparametric Tests

In a highly cogent article, Anderson (1961) compared parametric and nonparametric tests both with respect to the statistical considerations of significance level, power, and versatility and with respect to a theoretical consideration of the properties of the measurement scale, which in his paper was limited to numerical data in the dependent variable. Although in his detailed argument Anderson conceded that relative to both significance level and power of tests the difference between parametric and rank order data is not great, he emphasized that only the parametric tests afford the necessary versatility to fulfill the day-to-day requirements of psychological research and concluded that parametric methods are the basic tools in psychological statistics. From the standpoint of the theory of measurement, Anderson stressed that the presence of an interval scale is not a requirement for the formulation of a statistical inference from results furnished by application of a parametric test, argued convincingly that numerous practical features of invariance, or generality of findings, far overshadow refinements of scale typology, and pointed out that inferences regarding psychological meaning are actually primarily substantive and judgmental in focus rather than formal in the sense of being related to the reliability of data.

Robustness of the *t*-Test and the Wilcoxon-Mann-Whitney *U*-Test

The robustness of tests has been receiving an increasing amount of attention. Indeed, the extent to which a statistical test can retain its power as well as its nominal level of significance when assumptions underlying its application are violated is an important concern.

In terms of both a review of the literature and an empirical selection of samples from populations constructed to have specific characteristics, Boneau (1962) made a comparison of the power of the *t*-test against the power of the frequently used nonparametric Wilcoxon-Mann-Whitney *U*-test. For both one-tail and two-tail tests, various combinations of sample size and of population variances were taken systematically along with prescribed differences between population means being present. In general, when the variances of the two populations were homogeneous there was a remarkable degree of closeness in the empirically derived power curves of the *t*- and *U*-tests relative to samples of sizes 5 and 5 and of 15 and 15. In the instance of samples, one of size 5 and one of size 15, the degree of resemblance was still close, although the power of the *t*-test was slightly superior to that of the *U*-test. For the situation of normal distributions, of heterogenous variances in the two populations (the ratio of the variances being 4 to 1), and of samples of equal size, the power of both the *t*- and *U*-tests remained relatively high, although that of the *t*-test was slightly greater than that of the *U*-test.

When the size of one sample was three times larger than the other, and when the corresponding variance in the population from which the first sample was taken was also the larger of the two population variances, there was a substantial loss in power for both the *U*- and *t*-tests, although the power of *U* was affected less than was that of *t*. The empirical values for the significance levels α were reduced below the intended, or nominal, significance levels of .05 and .01—the departures being greater for *t* than for *U*. When the size of one sample was three times larger than the other, but the corresponding variance in the population from which the sample was chosen was the smaller of the two population variances, the effects were in the opposite direction. For both *U* and *t*, the empirical values for α were larger than the intended, or nominal, levels of .05 and .01—the discrepancies being larger for *t* than for *U*. Moreover, the power of *t* was greater than that of *U*. These results held for both one-tail and two-tail tests, although the two-tail tests did show slightly smaller discrepancies in distance between empirical power functions of *U* and *t* than did the one-tail tests. In short, when sample sizes differed and variances in the two populations were heterogeneous, the extent of distortion in the empirical alpha levels and in the degree of empirical power realized was relatively greater for *t* than for *U* for both one-tail and two-tail tests, although both tests were affected adversely.

In the instance of sampling from nonnormal populations (rectangular and exponential), it was noted that when the two distributions have the same shape, the *U*-test appeared to be consistently slightly more powerful than the *t*-test for small differences between population means, but not for larger differences. The shapes of the power curves were extremely similar. When one distribution was normal and the other was exponential, Boneau found that the empirically derived values for α were in general slightly closer to the nominal levels of α for *U* than for *t* with respect to both one-tail and two-tail tests and that the empirical power functions for *U* in general were only slightly higher than those for *t*. When the sizes of samples from two populations of different shape were increased, the slight advantage, if any, which U held over t disappeared. The author also suggested that, in general, empirically derived power functions for the t-test are probably much closer to the theoretical power function than are the seemingly more variable empirical curves for the U-test.

Considering his data, Boneau took a moderate view concerning whether the t-test or the U-test should be employed when assumptions underlying the t-test are violated, for both tests are subject to a certain amount of error. However, he recommended that one not desist from using the U-test in place of the t-test simply because it is less powerful than the t-test.

Other articles were concerned with the robustness of the t-test compared with that of the Wilcoxon statistic. Van der Vaart (1961) reviewed several studies and noted that when the variance of one population is zero—an extreme case—the relationship between sample sizes serves to determine which one of the two tests is to be preferred. In a comprehensive paper, Wetherill (1960) concluded that if two distributions are nearly normal but differ in shape, the Wilcoxon test, although sometimes slightly better than the t-test, is frequently inferior to it. Although some of his findings were similar to those noted by Boneau, Wetherill emphasized the complexity of the power function of the Wilcoxon-Mann-Whitney U statistic, which itself is highly dependent upon the mean, skewness, and kurtosis of the distribution. In his critical review of the U- and t-tests, Lubin (1962) described in detail the complexities of the problem from both the statistical and the psychological points of view. His cogent discussion is worth serious study and careful reflection, although there appears to be a possible contradiction in two adjacent sentences in one paragraph (p. 346, ll. 24-27).

Of considerable help to those who make frequent use of the Wilcoxon rank-sum test statistic in the two-sample case, new tables of the exact critical values have been developed (Verdooren, 1963). The lower significance levels of .001, .005, .01, .025, .05, and .10 are presented for combinations of sample size up to 25, along with two approximations and adjustments for ties.

Miscellaneous Contributions to Problems of Robustness, Power, and Efficiency

Reflecting a number of different approaches and points of view, several articles appeared that were concerned with problems of robustness, power, and efficiency of significance tests. In an extensive paper, Box and Tiao (1962) studied the problem of robustness through injecting into several models subjectively determined prior probability distributions that seemed plausible. They expressed confidence in the general applicability of such procedures. Through use of power functions and related criteria, David and Perez (1960) compared so-called quick tests of location and dispersion, which are frequently employed because of their computational simplicity and alleged robustness, with standard significance tests in order to ascertain the extent of similarity of conclusions regarding significance. Among several examples chosen for illustrative purposes, including an evaluation of the sign test relative to the *t*-test for paired data, the probability levels for the quick tests were only slightly larger than the chosen *a* levels for the standard tests of significance. Having noticed that the Kolmogorov-Smirnov test often has not yielded significant results for samples of moderate size when the chi-square test has done so, Durbin (1961) presented empirical results and then offered some methods of constructing exact tests of improved sensitivity that are intended to detect differences in two distributions.

In other papers concerned wth the efficiency of tests, Mikulski (1963) compared the efficiency of nonparametric procedures in the two-sample case relative to parametric ones when the true distribution diverges from the assumed one; Chanda (1963) determined the efficiency of nonparametric tests for discrete distributions against the most powerful tests to be found for these distributions when they are members of parametric families and suggested that the values of the relative efficiencies may be dependent upon the extent of the discreteness of the distribution; Capon (1961), in a highly abstract article, studied the asymptotic efficiency of several "locally most powerful rank tests"; and Sukhatme (1960) discussed the power of several nonparametric tests of dispersion.

New Tests of Significance

In the parametric case, two papers were concerned with the test of the hypothesis of equal means depending upon whether the variances in the normal populations are homogeneous or heterogeneous. To test the hypothesis that the means of two normal populations are equal when samples of small size are employed, McCullough, Gurland, and Rosenberg (1960) proposed two approaches for controlling the size of a test—that is, the value of the significance level *a*. In what they termed a *unilateral* case, a means was offered for controlling *a* when the variance of one specified population is known a priori to be larger than the other. Similarly, for what was called a *bilateral* case, in which no a priori information about population variances is available, a procedure was also suggested. Subsequently, Gurland and McCullough (1962) considered several approaches through which preliminary tests are carried out to determine whether the variances of two populations may be regarded as equal. In light of the results obtained, one proceeds to test the hypothesis that the two means are equal. The choice of level of significance of the preliminary test is anchored to the desired level of significance of the combined test. Detailed tables of size and power are furnished to guide one in studying the behavior of the proposed tests.

In the nonparametric domain, several papers have been concerned with the development of tests for differences in spread or dispersion. For two independent samples, Siegel and Tukey (1960) developed a nonparametric sum of ranks procedure to test whether two samples arise from the same population with respect to spread, or variability, in scores. In addition to inclusion of extensive tables of critical values for combinations of sample size from 1 to 20 ($n_1 \leqq n_2 < 20$), large sample procedures were furnished which also allow for a correction for ties. Since the *F*-test for differences in variance is relatively sensitive to departures from normality, this nonparametric test, which is entirely distribution-free in terms of customary randomization procedures, should be an important tool. Although this test is relatively sensitive to differences in dispersion when means or medians of the population are equal or nearly equal, its power decreases fairly rapidly as the difference between means increases and actually reaches zero when the populations do not overlap.

A related contribution by Ansari and Bradley (1960) was that of a rank-sum test for significance of difference in dispersion of two independent samples that arise from populations with continuous cumulative distribution functions. In their test, knowledge of the difference between medians of the two populations is a requirement. As mentioned previously, Sukhatme (1960) undertook a critical evaluation of the power properties of a number of nonparametric tests concerned with differences in dispersion of two populations and pointed out several deficiencies of which one should be aware.

In the instance of several samples, Bhapkar (1961) devised a new nonparametric procedure to test the hypothesis of equality of location (median). He discussed the asymptotic power and asymptotic efficiency of his test in relation to other well-known ones that have been proposed.

Additional References: Adcock (1960); Bhapkar (1961); Chapanis (1962).

Contingency Tables

Since July 1960, relatively little work appeared that involved the use of chi-square functions as compared with the amount of effort devoted during previous review periods. Much of what was reported regarding the use of chi square was in conjunction with the interpretation of data in contingency tables. In fact, considerable effort was directed toward problems of the analysis of data in contingency tables.

To assist the research worker confronted with those practical problems of interpreting multiway contingency tables which the general textbook does not answer, Lewis (1962) presented a general review of the principal methods of analysis along with a selection of computationally feasible procedures, which may be modified for use with different sampling plans as well as with theoretical parameters rather than with estimated ones.

Subsequently, in a comprehensive paper concerned with methods of comparing contingency tables, Goodman (1963) showed that the standard error formulas furnished by Lewis as well as by others for Stouffer's method of comparing 2x2 tables fail to yield consistent estimates. Through making certain modifications in these formulas, Goodman developed consistent estimators, pointed to substantial differences in the size of standard errors associated with his formulas and with those published previously, and offered generalizations of methods of comparing several 2x2 tables.

For the situation in which the 2x2 or 3x3 parent distribution is bivariate normal, Tallis (1962) furnished a means for the maximum likelihood estimation of correlation from contingency tables. In the instance of 2x2 tables in which the marginal totals are not fixed, Edwards (1963) showed that a measure of association should logically be functionally related to the crossratio. In applying a simple method to subdivide a large contingency table into a series of 2x2 tables, Daly (1962) obtained chi-square values for each of these tables—values that he concluded would furnish a feasible basis for a test of trend in most practical situations.

Two papers appeared that were concerned with errors of classification of data placed into cells of tables. Making use of four extreme situations of classification in which two groups were being compared in terms of their respective morbidity rates, Rogot (1961) noted that, relative to the differences for a set of error-free data, the observed difference in each of the error situations will be smaller and concluded that the presence of measurement error will serve to reduce the probability of detection of a genuine difference. When the parameter of interest for an alternative hypothesis is the difference between two true binominal probabilities that occur because of opportunities for misclassification, Walsh (1963) evaluated the efficiencies of tests for the circumstance in which the method of classification is subject to random errors and for the situation in which no classification errors take place and then concluded that the proposed tests are valid for both events.

Two other noteworthy papers were those of Bennett and Hsu (1960), who furnished power contour curves of the exact test for a 2x2 contingency table, and of Kincaid (1962) who prescribed a method involving a chi-square expression for combining a set of contingency tables each consisting of two rows and an arbitrary (but fixed) number *m* of columns.

Multiple Comparisons

Although relative to the control of error rate in the performance of multiple significance tests within the same experiment, Norton has described succinctly in Chapter III several of the major contributions to be found in the 1961 article by Duncan, in the 1961 and 1962 papers by

Ryan, and in the 1962 note by Wilson, attention also should be called to the critical discussion by Lubin (1962). In addition, four other papers were noteworthy. If prior to his investigation an experimenter can limit himself to a relatively small number of linear contrasts among a certain number of means, Dunn (1961) has demonstrated that the widths of the confidence intervals that can be established are shorter than those found by the familiar procedures of Scheffé and Tukey. After describing several factors that influence the number of significant findings in multiple significance tests, Block (1960) proposed a general method which he maintained allows for an unbiased inference of a group of findings considered as a set and then suggested as well as illustrated a high-speed computer-based approach to be used in the generation of an empirical sampling distribution appropriate to a particular data matrix. Harter (1960a) furnished improved estimates of critical values for the now well-known and popular Duncan's New Multiple Range Test, and Steel (1961) generalized Wilcoxon's rank-sum test in order that multiple comparisons for each of several random samples of equal size given experimental treatments could be made against a control group.

Concluding Remarks

As a participant in the preparation of the chapters upon statistical methodology for the December 1957 and December 1960 issues of the REVIEW, this writer has noted what appear to be a few changes in emphasis in the statistical literature from July 1, 1954, to July 1, 1963. During the period from 1954 to 1957, considerable stress was placed upon development of nonparametric methods, perhaps at the expense of effort being directed toward augmenting classical parametric theory. From 1957 to 1960, the amount of attention given to nonparametric statistics appeared to decrease, and renewed interest in the development of parametric procedures was evident.

The most recent three-year period, which seems to have been one of consolidation and conciliation, has been marked by attempts to furnish a unified view of statistical inference and statistical analysis, as evidenced by the general and inclusive expository presentations of Birnbaum (1961, 1962), Pearson (1962), and Tukey (1962) as well as by the somewhat focalized but relatively comprehensive contributions of Brillinger (1962), Fraser (1961), and Savage and others (1962). Evidence of a partial *rapprochement* between nonparametric and parametric methods was apparent in the papers by Anderson (1961) and Boneau (1962) and in others in which parametric and nonparametric procedures were compared. Attention in several articles has been given to problems of the comparative power and efficiency of nonparametric and parametric tests. The most noteworthy new direction is that furnished by Bayesian statistics, the application of which affords an exciting and promising approach in research both to the design of educational and psychological investigations and to the analysis and evaluation of the forthcoming data.

Bibliography

ALEXANDER, HOWARD W. *Elements of Mathematical Statistics.* New York: John Wiley & Sons, 1961. 367 pp.

ANDERSON, NORMAN H. "Scales and Statistics: Parametric and Nonparametric." *Psychological Bulletin* 58: 305-16; July 1961.

ANSARI, A. R., and BRADLEY, R. A. "Rank-Sum Tests for Dispersions." *Annals of Mathematical Statistics* 31: 1174-89; December 1960.

BENNETT, B. M., and HSU, P. "On the Power Function of the Exact Test for the 2x2 Contingency Table." *Biometrika* 47: 393-98; December 1960.

BHAPKAR, V. P. "A Nonparametric Test for the Problem of Several Samples." *Annals of Mathematical Statistics* 32: 1108-17; December 1961.

BINDER, ARNOLD. "Further Considerations on Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 70: 107-15; January 1963.

BIRNBAUM, ALLAN. "A Unified Theory of Estimation, I." *Annals of Mathematical Statistics* 32: 112-35; March 1961.

BIRNBAUM, ALLAN. "On the Foundations of Statistical Inference." *Journal of the American Statistical Association* 57: 269-326; June 1962.

BLOCK, JACK. "On the Number of Significant Findings To Be Expected by Chance." *Psychometrika* 25: 369-80; December 1960.

BLYTH, COLIN R., and HUTCHINSON, DAVID W. "Table of Neyman-Shortest Unbiased Confidence Intervals for the Binomial Parameter." *Biometrika* 47: 381-91; December 1960.

BONEAU, C. ALAN. "A Comparison of the Power of the U and t Tests." *Psychological Review* 69: 246-56; May 1962.

BOX, G. E. P., and TIAO, G. C. "A Further Look at Robustness via Bayes' Theorem." *Biometrika* 49: 419-32; December 1962.

BRILLINGER, DAVID R. "Examples Bearing on the Definition of Fiducial Probability with a Bibliography." *Annals of Mathematical Statistics* 33: 1349-55; December 1962.

CAPON, JACK. "Asymptotic Efficiency of Certain Locally Most Powerful Rank Tests." *Annals of Mathematical Statistics* 32: 88-100; March 1961.

CHANDA, K. C. "On the Efficiency of Two-Sample Mann-Whitney Test for Discrete Populations." *Annals of Mathematical Statistics* 34: 612-17; June 1963.

DALY, C. "A Simple Test for Trends in a Contingency Table." *Biometrics* 18: 114-19; March 1962.

DAVID, H. A., and PEREZ, CARMEN A. "On Comparing Different Tests of the Same Hypothesis." *Biometrika* 47: 297-306; December 1960.

DUNN, OLIVE JEAN. "Multiple Comparisons Among Means." *Journal of the American Statistical Association* 56: 52-64; March 1961.

DURBIN, J. "Some Methods of Constructing Exact Tests." *Biometrika* 48: 41-55; June 1961.

EDWARDS, A. W. F. "The Measure of Association in a 2x2 Table." *Journal of the Royal Statistical Society* (Series A, General) 126: 109-14; No. 1, 1963.

EDWARDS, WARD; LINDMAN, HAROLD; and SAVAGE, LEONARD J. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70: 193-242; May 1963.

FIELDS, R. I.; KRAMER, C. Y.; and CLUNIES-ROSS, C. W. "Joint Estimation of the Parameters of Two Normal Populations." *Journal of the American Statistical Association* 57: 446-54; June 1962.

FRASER, D. A. S. "On Fiducial Inference." *Annals of Mathematical Statistics* 32: 661-76; September 1961.

FREUND, JOHN E. *Modern Elementary Statistics.* Second edition. Englewood Cliffs, N.J.: Prentice-Hall, 1960. 413 pp.

FREUND, JOHN E. *Mathematical Statistics.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 390 pp.

GAGE, N. L., editor. *Handbook of Research on Teaching.* Chicago: Rand McNally & Co., 1963. 1218 pp.

GOODMAN, LEO A. "On Methods for Comparing Contingency Tables." *Journal of the Royal Statistical Society* (Series A, General) 126: 94-108; No. 1, 1963.

GRANT, DAVID A. "Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 69: 54-61; January 1962.

GRAYBILL, FRANKLIN A., and MORRISON, ROBERT D. "Sample Size for a Specified Width Confidence Interval on the Variance of a Normal Distribution." *Biometrics* 16: 636-41; December 1960.

GURLAND, JOHN, and MCCULLOUGH, ROGER S. "Testing Equality of Means After a Preliminary Test of Equality of Variances." *Biometrika* 49: 403-17; December 1962.

HALPERIN, MAX. "Confidence Intervals from Censored Samples." *Annals of Mathematical Statistics* 32: 828-37; September 1961.

HAMMOND, KENNETH R., and HOUSEHOLDER, JAMES E. *Introduction to the Statistical Method.* New York: Alfred A. Knopf, 1962. 417 pp.

HARRIS, CHESTER W. "Statistical Methods." *Encyclopedia of Educational Research.* Revised edition. (Edited by Chester W. Harris.) New York: Macmillan Co., 1960. pp. 1397-1410.

HARTER, H. LEON. "Critical Values for Duncan's New Multiple Range Test." *Biometrics* 16: 671-85; December 1960. (a)

HARTER, H. LEON. "Tables of Range and Studentized Range." *Annals of Mathematical Statistics* 31: 1122-47; December 1960. (b)

HODGES, J. L., JR., and LEHMANN, E. L. "Estimates of Location Based on Rank Tests." *Annals of Mathematical Statistics* 34: 598-611; June 1963.

HORST, PAUL. *Matrix Algebra for Social Scientists.* New York: Holt, Rinehart and Winston, 1963. 517 pp.

JEFFREYS, HAROLD. *Theory of Probability.* Third edition. New York: Oxford University Press, 1961. 447 pp.

KENDALL, MAURICE G., and STUART, ALAN. *The Advanced Theory of Statistics. Vol. 2: Inference and Relationship.* Revised edition. New York: Hafner Publishing Co., 1961. 676 pp.

KINCAID, W. M. "The Combination of 2x*m* Contingency Tables." *Biometrics* 18: 224-28; June 1962.

KOZELKA, ROBERT M. *Elements of Statistical Inference.* Reading, Mass.: Addison-Wesley Publishing Co., 1961. 150 pp.

LEWIS, B. N. "On the Analysis of Interaction in Multi-Dimensional Contingency Tables." *Journal of the Royal Statistical Society* (Series A, General) 125: 88-117; No. 1, 1962.

LUBIN, ARDIE. "Statistics." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1962. Vol. 13, pp. 345-70.

MCCULLOUGH, ROGER S.; GURLAND, JOHN; and ROSENBERG, LLOYD. "Small Sample Behaviour of Certain Tests of the Hypothesis of Equal Means Under Variance Heterogeneity." *Biometrika* 47: 345-53; December 1960.

MCNEMAR, QUINN. *Psychological Statistics.* Third edition. New York: John Wiley & Sons, 1962. 451 pp.

MADANSKY, ALBERT. "More on Length of Confidence Intervals." *Journal of the American Statistical Association* 57: 586-89; September 1962.

MIKULSKI, PIOTR WITOLD. "On the Efficiency of Optimal Nonparametric Procedures in the Two Sample Case." *Annals of Mathematical Statistics* 34: 22-32; March 1963.

MOSTELLER, FREDERICK, and WALLACE, DAVID L. "Inference in an Authorship Problem." *Journal of the American Statistical Association* 58: 275-309; June 1963.

MOSTELLER, FREDERICK; ROURKE, ROBERT E. K.; and THOMAS, GEORGE B., JR. *Probability with Statistical Applications.* Reading, Mass.: Addison-Wesley Publishing Co., 1961. 478 pp.

NUNNALLY, JUM. "The Place of Statistics in Psychology." *Educational and Psychological Measurement* 20: 641-50; Winter 1960.

PACHARES, JAMES. "Tables of Confidence Limits for the Binomial Distribution." *Journal of the American Statistical Association* 55: 521-33; September 1960.

PEARSON, E. S. "Some Thoughts on Statistical Inference." *Annals of Mathematical Statistics* 33: 394-403; June 1962.

PEATMAN, JOHN G. *Introduction to Applied Statistics.* New York: Harper & Row, 1963. 458 pp.

PRATT, JOHN W. "Length of Confidence Intervals." *Journal of the American Statistical Association* 56: 549-67; September 1961.

Rogot, Eugene. "A Note on Measurement Errors and Detecting Real Differences." *Journal of the American Statistical Association* 56: 314-19; June 1961.

Rozeboom, William W. "The Fallacy of the Null-Hypothesis Significance Test." *Psychological Bulletin* 57: 416-28; September 1960.

Savage, Leonard J., and others. *The Foundations of Statistical Inference.* New York: John Wiley & Sons, 1962. 112 pp.

Siegel, Sidney, and Tukey, John W. "A Nonparametric Sum of Ranks Procedure for Relative Spread in Unpaired Samples." *Journal of the American Statistical Association* 55: 429-45; September 1960.

Steel, Robert G. D. "Some Rank Sum Multiple Comparisons Tests." *Biometrics* 17: 539-52; December 1961.

Sukhatme, Balkrishna V. "Power of Some Two-Sample Non-Parametric Tests." *Biometrika* 47: 355-62; December 1960.

Tallis, G. M. "The Maximum Likelihood Estimation of Correlation from Contingency Tables." *Biometrics* 18: 342-53; September 1962.

Tatsuoka, Maurice M., and Tiedeman, David V. "Statistics as an Aspect of Scientific Method in Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 4, pp. 142-70.

Tukey, John W. "The Future of Data Analysis." *Annals of Mathematical Statistics* 33: 1-67; March 1962.

Van der Vaart, H. R. "On the Robustness of Wilcoxon's Two Sample Test." *Biometrics* 17: 183; March 1961.

Verdooren, L. R. "Extended Tables of Critical Values for Wilcoxon's Test Statistic." *Biometrika* 50: 177-86; June 1963.

Walsh, John E. *Handbook of Nonparametric Statistics.* Princeton, N.J.: D. Van Nostrand Co., 1962. 549 pp.

Walsh, John E. "Loss in Test Efficiency Due to Misclassification for 2x2 Tables." *Biometrics* 19: 158-62; March 1963.

Wetherill, G. B. "The Wilcoxon Test and Non-Null Hypotheses." *Journal of the Royal Statistical Society* (Series B, Methodological) 22: 402-18; No. 2, 1960.

Wilks, Samuel S. *Mathematical Statistics.* New York: John Wiley & Sons, 1962. 644 pp.

Additional References

Adcock, C. J. "A Note on Combining Probabilities." *Psychometrika* 25: 303-305; September 1960.

Bhapkar, V. P. "Some Nonparametric Median Procedures." *Annals of Mathematical Statistics* 32: 846-63; September 1961.

Birnbaum, Allan. "Confidence Curves: An Omnibus Technique for Estimation and Testing Statistical Hypotheses." *Journal of the American Statistical Association* 56: 246-49; June 1961.

Chapanis, Alphonse. "An Exact Multinomial One-Sample Test of Significance." *Psychological Bulletin* 59: 306-10; July 1962.

Harter, H. Leon. "The Use of Sample Ranges in Setting Exact Confidence Bounds for the Standard Deviation of a Rectangular Population." *Journal of the American Statistical Association* 56: 601-609; September 1961.

Harter, H. Leon. "Percentage Points of the Ratio of Two Ranges and Power of the Associated Test." *Biometrika* 50: 187-94; June 1963.

Murthy, M. N. "Variance and Confidence Interval Estimation." *Sankhyā* (Series B) 24: 1-12; January 1962.

Pachares, James. "Tables for Unbiased Tests on the Variance of a Normal Population." *Annals of Mathematical Statistics* 32: 84-87; March 1961.

Sarhan, Ahmed E., and Greenberg, Bernard G. *Contributions to Order Statistics.* New York: John Wiley & Sons, 1962. 482 pp.

Savage, I. Richard. *Bibliography of Nonparametric Statistics.* Cambridge, Mass.: Harvard University Press, 1962. 284 pp.

CHAPTER III

Developments in Analysis of Variance and Design of Experiments

DEE W. NORTON

THE PRESENT chapter was prepared with two groups of readers in mind. The first—and smaller—group includes those persons whose primary concern is the development and interpretation of research methods applicable to studies in education and the social sciences. The second group is comprised of those who are planning and analyzing substantive research studies in these areas—the content researchers. Both types of readers will find it essential to refer to the cited sources before drawing firm conclusions regarding the worth of the work assayed so briefly in the following paragraphs. In doing so, the content researchers will often find it necessary to seek assistance from a member of the first group of readers.

Contrary to the practice of Stanley and of Collier and Meyer in their coverage of design and analysis in the December 1957 REVIEW and the December 1960 REVIEW, respectively, this chapter does not necessarily omit references cited in the chapters on statistics in the *Annual Review of Psychology*. In 1960, the *Annual Review* changed to an alternate-year coverage of the broad topic of statistical theory; hence, the overlap between this chapter and Lubin's (1962) contribution is not pronounced, and such overlap as has occurred does not seem undesirable.

Philosophical Basis for Inference from Experimental Studies

In a long and penetrating article, Birnbaum (1962) developed an argument which may have significant impact on the design, analysis, and interpretation of experimental studies. Birnbaum considered the interrelated problems of informative inference, experimental evidence, and statistical evidence. After noting the widespread acceptance of the principles of sufficiency and conditionality, Birnbaum showed that these two principles taken together are mathematically equivalent to the likelihood principle. Taking issue with the usual interpretation of experimental outcomes in terms of relative frequency in conceptual replications, he argued that the evidential meaning of any outcome is characterized completely by the likelihood function, without reference to the structure of the experiment. He further argued that any mode of evidential interpretation which would lead to different conclusions from sequential or nonsequential experiments is inconsistent with the likelihood principle. Appearances to the contrary, Birnbaum indicated that he was interested in an approach to informative inference that did not depend on the Bayesian principle of inverse probability.

Birnbaum stated at one point that his findings would have radical consequences for everyday practice. Yet he later seemed to imply that conventional procedures would be adequate if thorough consideration were given in the planning of an experiment to control both types of error for a specified type and degree of falsity of the hypothesis under test.

The issues explored quite formally by Birnbaum had earlier been touched by Rozeboom (1960). Rozeboom criticized the widespread obeisance to .05 significance levels and urged greater use of the likelihood function, confidence intervals, and inverse probabilities. He did not, in this reviewer's opinion, give adequate consideration to those situations in which an experiment is carefully planned with formal attention to power and sensitivity.

In a somewhat different vein, Tukey (1962) reviewed at length the current status of analysis of variance and similar techniques. He emphasized the need for (a) more efficient techniques to deal with "wild" data, (b) further investigation of the potentialities of half-normal plotting to identify error sources, and (c) additional work on the transformation problem.

Additional References: Albert (1961); Neyman and Scott (1960); Pearson (1962).

Mixed Model—Repeated Measurement Designs

One Fixed Factor—Repeated Measures on the Random Factor

Gaito (1961) presented a worthwhile analysis of the difficulties involved with designs in which each subject undergoes each of a fixed set of treatments. He presented expected mean squares for cases in which order of administration is (a) constant for all subjects, (b) randomized for each subject, and (c) counterbalanced with and without modified analysis. His emphasis on the fact that counterbalancing does not get rid of practice effects was well placed. In considering the cross-over arrangement in which simple order becomes an additional factor and sequence effects appear as treatment-order interaction, Gaito did not present an entirely convincing defense for his claim of negative bias in the F-test of the within-subjects component of TxO. If there were no treatment effect, it seems clear that there would be order-subject interaction effects within each group and that these effects would consequently produce an apparent order-group interaction. If such an argument is extended, it seems reasonable to expect in MS_{TxO} all the same sources of variation as are in the within-subjects residual term. This argument is important, since if a justifiable test of treatment-order interaction is available, an experimenter can determine whether the attempt to increase precision by repeated observation under counterbalanced orders has been successful. If a sequence effect is identified, it is only necessary, as Gaito clearly pointed out, to resort to the simple randomized analysis of the data observed on the first administration only.

Imhof (1962) questioned the validity of the usual test of no main effect of treatments in treatments-by-subjects design. As Box and other writers have pointed out in the past few years, Imhof noted that the ratio MS_T/MS_{TS} does not (except for the two-treatment case) follow the F distribution with (t-1) and (t-1)(s-1) degrees-of-freedom unless the within- and between-treatment dispersion matrix has diagonal elements σ^2 and off-diagonal elements $\rho\sigma^2$. Imhof went on to show that for small degrees-of-freedom the actual risk of a Type I error in a nominal .01-level test of the hypothesis of equal treatment-population means ranged from .01 to .08, depending on the pattern of disparity among the nonzero characteristic roots of the dispersion matrix. He also provided tables for an F-test with modified degrees of freedom, the modification depending on a test of the interaction mean square when multiple measures per subject per treatment make an error term for such a test available. Imhof did not compare his modified test with that advanced earlier by Box, but did find it rather more powerful with small degrees-of-freedom than the alternative test based on the Hotelling T^2, which does not place restrictions on the form of the dispersion matrix.

Additional References: McHugh and Wall (1962); Roy and Cobb (1960).

Two Fixed Factors—Repeated Measures on a Random Factor Under One of the Fixed Factors

There is no brief title which adequately characterizes designs of the type considered here. In some quarters they are known as partial hierarchical designs, in others as Lindquist Type I. In any event, they are widely used in educational and psychological research studies. Danford, Hughes, and McNee (1960) provided an exposition of the assumptions underlying this design which will be of particular interest to practicing research workers. The authors showed how to check on the assumptions of equality of variances and equality of covariances within each of b levels of the between-groups factor and for equality of the b dispersion matrices. They also illustrated the use of Hotelling's T^2 statistic when the several dispersion matrices are the same but not symmetrical in form. They called attention to the important fact that when the repeated measures are on successive trials, the correlations between adjacent trials will often be greater than between separated trials—a fact implying violation of the assumptions. They also suggested an analysis for testing for parallelism of linear trends among the b treatment groups, but failed to emphasize that this test involves the implicit assumption that a linear trend is an adequate representation of the relationship in each b condition.

Complex Mixed Designs Involving Repeated Measurements

Imhof (1960) explored a type of design which seems likely to have considerable applicability to studies in education and the behavioral sciences but which has received little attention in the research literature. The design involves one fixed factor and two crossed random factors: e.g., each of a random sample of s subjects responds to each of a random sample of i stimuli under each of t fixed treatments. Imhof showed that despite appearances from the table of expected mean squares, the ratio of the mean square for interaction between the fixed factor and one random factor to the mean square for triple interaction is distributed as F only when the conditions specified for a one-fixed—one-random factor design are satisfied in three distinct senses. Furthermore, he noted that no straightforward test is generally available for the main effect of the fixed factor.

Stanley (1961b) discussed a similar type of design as applied to a reliability study involving subjects, raters, and traits. He pointed out the implications for estimation and tests of significance of various decisions about treating the several factors as fixed or random and showed how to secure estimates of selected reliability coefficients. In a later article, Stanley (1961a) exhibited an analysis of a very complex design with crossed random factors (subjects and raters) and multiple measures per subject-rater combination. In addition, the raters and the subjects were each subdivided on classificatory bases. Thus, within each type-of-rater by type-of-subject cell, the design was of the sort considered in his earlier article. Stanley showed how to form synthetic error terms for approximate tests of the main effects of the classificatory factors.

Additional References: Graybill and Hultquist (1961); Harter (1961); Thompson (1962).

Multifactor Designs Without Repeated Measurements

Randomized Block Designs

Collier (1961) made a very important contribution to the extensive literature on designs in which a stratifying factor is introduced to increase the precision of what might otherwise be a completely randomized one-factor design. He gave special attention to the effect of treatment-subject interaction on expected mean squares and choice of error terms in designs in which each of the several blocks is comprised of subjects presumably homogeneous on some variable related to the criterion measure. Within each block (level), random subgroups of subjects are assigned to each treatment. Making use of finite randomization models, Collier emphasized the need to conceptualize a mean score for each subject over all treatments, even though each subject is in fact measured only under the particular treatment resulting from random assignment. He showed that in terms of expected mean squares, proper error terms exist for the important case of fixed treatments, fixed blocks, and random subjects. In view of the stringent assumptions for exact F distributions in repeated measurement designs in which treatment-subject interaction effects are explicitly observed—*see* Imhof (1962)—one might have wished for more elaboration of Collier's comment that the adequacy of his tests depends on how well the empirical distribution of the mean square ratio under all possible randomizations follows the presumed F distribution.

The issue just noted was explored to some extent by Folks and Kempthorne (1960) in connection with the efficiency of blocking in incomplete block designs, i.e., designs in which less than all treatments occur in each block. Incomplete designs have not yet been widely used in educational and psychological research, probably because they inevitably involve some confounding of effects and, hence, assumptions about nonexistence of certain effects. Their wide use in industrial and agricultural research and their potential applicability to investigatory studies in clinical areas suggests that incomplete designs should receive more attention.

Additional References: Addelman (1963); Potthoff (1963).

Response Surface and Related Designs

Response surfaces may be crudely characterized as generalizations in several dimensions of the familiar two-dimensional interaction patterns. In investigating response patterns that are a function of many interrelated variables, the problem arises of choosing a design matrix that specifies the levels at which each of the factors should be represented. A key article on this issue was that by Box and Behnken (1960). They were specifically concerned with optimal procedures for choosing a design matrix that permits the fitting of a second-degree polynomial to the response surface. They developed a method for producing rotatable designs in which the error variance of the response estimates is constant for all points (treatment-combinations) equidistant from the center of the factor space. Bose and Carter (1962) presented a model which permits estimation of the response corresponding to any point in the factor space, not just at those points represented in the design—an estimation which does not require equal spacing of the levels of the various factors.

Dempster (1961) continued his theoretical investigation of the properties of multifactor incomplete designs formed by selecting treatment-combinations at random from the set of all possible combinations. Wilk, Gnanadesikan, and Freeny (1963) developed a method of estimating error variance in multifactor designs from the M smallest squared single degree-of-freedom contrasts out of K such contrasts available. In limited empirical checks with M/K ratios from 0.03 to 1.0 and K's from 10 to 200, they found their estimate to be generally less than 5 percent too large. Their technique may prove quite useful in designs in which there is no effect assuredly free of treatment effects.

Additional Reference: Ehrenfeld and Zacks (1961).

Latin Square Designs

Sheehe and Bross (1961) presented an interesting method for constructing Latin square designs which are balanced with respect to (a) the number of treatments preceding every treatment, (b) the set of all predecessor treatments, and (c) the set of immediate predecessors. They did not, however, recommend any specific analysis of their designs which might take adequate account of the balancing. Ogilvie (1963) proposed a technique for removing time trend effects in a one-factor repeated measurement design employing balanced sets of Latin squares. The method involves the tenuous assumption of no treatment-subject interaction; hence, as Gaito (1961) and Collier (1961) pointed out, it is likely to be of limited applicability.

Special Topics in Design and Analysis

Multiple Comparisons and Separation of Populations

The problem of controlling error rate in making multiple tests of significance within the same experiment is not yet satisfactorily resolved. Starting with a detailed Bayesian analysis of the two-decision and three-decision t-test problems, Duncan (1961) developed a "minimum-average-risk" multiple decision rule for all possible two-treatment comparisons among n means in a one-factor completely randomized experiment. His rule requires specification of (a) a loss-ratio assumed constant for all comparisons, reflecting the relative seriousness of two-treatment Type I and Type II errors, (b) the variance of the (presumably normal) a priori distribution of the difference between population means, and (c) additive losses. In testing his rule, Duncan found that the critical difference for comparing means was almost independent of the total number of means involved—a circumstance that was due largely, he emphasized, to the condition of additive losses. His findings will be heartening to those who feel as does this reviewer that most multiple comparison procedures reflect excessive concern with the risk of an n-treatment Type I error in spite of the fact that two-treatment differences are commonly of primary interest.

Wilson (1962) made a good case for setting error rates on a per hypothesis basis, as has been the common though usually unacknowledged practice. He also emphasized the complete interdependence among the per hypothesis, per experiment, and experimentwise error rates and, hence, the impossibility of choosing among them on purely logical grounds. Ryan (1962) reiterated his preference for setting rates for the entire experiment, but admitted the complexity involved in attempting to balance the risks of the two types of error with respect to the over-all hypothesis and the many embedded hypotheses. In an earlier article, Ryan (1960) presented his generalized method for making pairwise comparisons on statistics other than the mean.

In order to keep an upper limit, say a, on the experimentwise error rate, the largest difference among n groups was tested at a nominal $2a/n(n\text{-}1)$ level. If significant, then the two next-largest differences were tested at a nominal $2a/n(n\text{-}2)$ level. This relaxing of significance levels for tests made within a significant over-all test occurs in several multiple comparison procedures. For example, this relaxation takes place when a significant F-test is followed with t-tests in place of the use of the mathematically more appropriate Scheffé S-test.

Lehmann (1961) considered the problem of selecting in a one-way design those treatments which produce a higher mean response than that associated with a standard treatment or with a standard set by the experimenter. He not only indicated why the standard treatment should in general be explicitly included in the experiment but also developed some fairly simple procedures that minimize the expected proportion of false-positive classifications.

Additional References: Bartholomew (1961); Fabian (1962); Hogg (1962).

Additivity

Additivity of effects is an important consideration in formulating an appropriate model for many analysis of variance designs. Mandel (1961) suggested a modification of Tukey's test for nonadditivity in a two-way design with one observation per cell. Mandel made the usual stringent assumptions that the true relationship in each row of the design is linear and that the intrinsic nonadditivity which may be involved is due entirely to differences among slopes from row to row. He did not, however, make the additional assumption that these differences in slope are directly dependent on main row effects. Elston (1961) proposed a method for determining whether nonadditivity of the Tukeyian sort can be removed by a certain type of monotone scale transformation. Lubin (1961) offered some useful comments about the implications of nonadditivity for interpretation of tests on main effects in factorial designs. He did not, however, distinguish carefully between the nonadditivity that affects *interpretations* in completely randomized designs and the nonadditivity that bears on the *validity* of F-tests in repeated measurement designs.

Covariance Analysis

An ingenious development pertaining to covariance was presented by Robson and Atkinson (1960). They showed how one may improve on the power of the usual test for homogeneity of linear regression by restricting the alternatives against which this hypothesis is tested. For a three-group situation they specified the following alternatives: (a) the three regression lines intersect at a common point with slopes proportional to intercept, and (b) the three regression lines intersect at different points with slopes a quadratic function of intercept. The among-regression-coefficients sum-of-squares was partitioned into two one-degree-of-freedom orthogonal components for regression on adjusted intercepts, and each component was tested against the pooled within-group deviation from individual regressions.

Additional References: Cox (1962); Curnow and Sharpe (1962).

Multivariate Analysis of Variance

Increasing interest has been shown in the problem of how to deal with designs in which multiple criterion measures are obtained on each experimental unit (e.g., subject). Smith, Gnanadesikan, and Hughes (1962) noted that in such situations the use of separate univariate analyses takes no account of the correlations among the response measures. They commented that a high over-all risk of Type I error may result. However, this criticism seems a bit casual in view of the difficulty of specifying precisely what hypotheses are of interest. These authors illustrated this very point by exhibiting three distinct (and complex) criteria of equality among treatment effects and by noting the arbitrary nature of the choice among them. Ito (1962) compared the power of the classical T^2- and W-tests which are available for dealing with multiple response measures. He tabled approximate powers of the two tests against selected classes of alternatives for an .05 significance level and found little difference for samples size of 100 to 200.

Additional Reference: Bhapkar (1960).

Assumptions in Two-Group Designs

McCullough, Gurland, and Rosenberg (1960) made an extensive investigation of ways to make use of knowledge about differences between population variances in testing the hypotheses of equal means in a two-group simple-randomized design. For the case in which one variance is known to be greater than the other, their proposed statistic compared favorably in power with the modified *t*-test appropriate when the variance ratio is known. For the case in which the variance ratio is completely unknown, the authors found it difficult to devise a statistic which would maintain a reasonably constant risk of Type I error. Gurland and McCullough (1962) extended this work to include preliminary tests on the sample variances and formulated explicit sequential procedures which have real promise.

Concluding Remarks

A truly significant contribution to the growing list of books on design and analysis was made by Winer (1962). He covered a great variety of designs with remarkable thoroughness. He emphasized the structural model underlying each design as well as the conditions that must be satisfied for valid statistical tests. His treatment of repeated measurement designs was particularly outstanding. Numerical illustrations or examples were provided for most designs, but interpretive discussion seemed rather limited in some instances. Winer also devoted more attention to complex incomplete designs than seems warranted by current needs in educational and psychological research. Edwards (1960) brought out a revision of his well-known text in design, with greater emphasis on individual comparisons, on trend analysis, and on graphical interpretation of interactions. Campbell and Stanley (1963) gave valuable specific criticisms of a variety of designs used in research on teaching. Graybill (1961) produced a highly technical but much-needed integration of analysis of variance and covariance techniques with a more generalized approach to the study of relationships.

Bibliography

Birnbaum, Allan. "On the Foundations of Statistical Inference." *Journal of the American Statistical Association* 57: 269-326; June 1962.

Bose, R. C., and Carter, R. L. "Response Model Coefficients and the Individual Degrees of Freedom of a Factorial Design." *Biometrics* 18: 160-71; June 1962.

Box, G. E. P., and Behnken, D. W. "Simplex-Sum Designs: A Class of Second Order Rotatable Designs Derivable from Those of First Order." *Annals of Mathematical Statistics* 31: 838-64; December 1960.

Campbell, Donald T., and Stanley, Julian C. "Experimental and Quasi-Experimental Designs for Research on Teaching." *Handbook of Research on Teaching.* (Edited by N. L. Gage.) Chicago: Rand McNally & Co., 1963. Chapter 5, pp. 171-246.

Collier, Raymond O., Jr. "The Effects of Subject-Treatment Interaction in a Randomized Block, One Factor Experiment." *Educational and Psychological Measurement* 21: 819-29; Winter 1961.

Danford, M. B.; Hughes, Harry M.; and McNee, R. C. "On the Analysis of Repeated-Measurements Experiments." *Biometrics* 16: 547-65; December 1960.

Dempster, A. P. "Random Allocation Designs II: Approximate Theory for Simple Random Allocation." *Annals of Mathematical Statistics* 32: 387-405; June 1961.

Duncan, David B. "Bayes Rules for a Common Multiple Comparisons Problem and Related Student-t Problems." *Annals of Mathematical Statistics* 32: 1013-33; December 1961.

Edwards, Allen L. *Experimental Design in Psychological Research.* Revised edition. New York: Holt, Rinehart and Winston, 1960. 389 pp.

Elston, R. C. "On Additivity in the Analysis of Variance." *Biometrics* 17: 209-19; June 1961.

Folks, John Leroy, and Kempthorne, Oscar. "The Efficiency of Blocking in Incomplete Block Designs." *Biometrika* 47: 273-83; December 1960.

Gaito, John. "Repeated Measurements Designs and Counterbalancing." *Psychological Bulletin* 58: 46-54; January 1961.

Graybill, Franklin A. *An Introduction to Linear Statistical Methods. Volume 1.* New York: McGraw-Hill Book Co., 1961. 463 pp.

Gurland, John, and McCullough, Roger S. "Testing Equality of Means After a Preliminary Test of Equality of Variances." *Biometrika* 49: 403-17; December 1962.

Imhof, J. P. "A Mixed Model for the Complete Three-Way Layout with Random-Effects Factors." *Annals of Mathematical Statistics* 31: 906-28; December 1960.

Imhof, J. P. "Testing the Hypothesis of No Fixed Main-Effects in Scheffé's Mixed Model." *Annals of Mathematical Statistics* 33: 1085-95; September 1962.

Ito, Koichi. "A Comparison of the Powers of Two Multivariate Analysis of Variance Tests." *Biometrika* 49: 455-62; December 1962.

Lehmann, E. L. "Some Model I Problems of Selection." *Annals of Mathematical Statistics* 32: 990-1011; December 1961.

Lubin, Ardie. "The Interpretation of Significant Interaction." *Educational and Psychological Measurement* 21: 807-17; Winter 1961.

Lubin, Ardie. "Statistics." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1962. Vol. 13, pp. 345-70.

McCullough, Roger S.; Gurland, John; and Rosenberg, Lloyd. "Small Sample Behaviour of Certain Tests of the Hypothesis of Equal Means Under Variance Heterogeneity." *Biometrika* 47: 345-53; December 1960.

Mandel, John. "Non-Additivity in Two-Way Analysis of Variance." *Journal of the American Statistical Association* 56: 878-88; December 1961.

Ogilvie, John C. "A Simple Method for the Elimination of Individual Trends in the Analysis of Balanced Sets of Latin Squares." *Biometrics* 19: 264-72; June 1963.

Robson, D. S., and Atkinson, G. F. "Individual Degrees of Freedom for Testing Homogeneity of Regression Coefficients in a One-Way Analysis of Covariance." *Biometrics* 16: 593-605; December 1960.

Rozeboom, William W. "The Fallacy of the Null-Hypothesis Significance Test." *Psychological Bulletin* 57: 416-28; September 1960.

Ryan, Thomas A. "Significance Tests for Multiple Comparison of Proportions, Variances, and Other Statistics." *Psychological Bulletin* 57: 318-28; July 1960.

Ryan, Thomas A. "The Experiment as the Unit for Computing Rates of Error." *Psychological Bulletin* 59: 301-305; July 1962.

Sheehe, Paul R., and Bross, Irwin, D. J. "Latin Squares To Balance Immediate Residual, and Other Order, Effects." *Biometrics* 17: 405-14; September 1961.

Smith, H.; Gnanadesikan, R.; and Hughes, J. B. "Multivariate Analysis of Variance (MANOVA)." *Biometrics* 18: 22-41; March 1962.

Stanley, Julian C. "Analysis of a Doubly Nested Design." *Educational and Psychological Measurement* 21: 831-37; Winter 1961. (a)

Stanley, Julian C. "Analysis of Unreplicated Three-Way Classifications, with Applications to Rater Bias and Trait Independence." *Psychometrika* 26: 205-19; June 1961. (b)

Tukey, John W. "The Future of Data Analysis." *Annals of Mathematical Statistics* 33: 1-67; March 1962.

Wilk, M. B.; Gnanadesikan, R.; and Freeny, Anne E. "Estimation of Error Variance from Smallest Ordered Contrasts." *Journal of the American Statistical Association* 58: 152-60; March 1963.

Wilson, Warner. "A Note on the Inconsistency Inherent in the Necessity To Perform Multiple Comparisons." *Psychological Bulletin* 59: 296-300; July 1962.

Winer, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill Book Co., 1962. 672 pp.

Additional References

Addelman, Sidney. "Techniques for Constructing Fractional Replicate Plans." *Journal of the American Statistical Association* 58: 45-71; March 1963.

Albert, Arthur E. "The Sequential Design of Experiments for Infinitely Many States of Nature." *Annals of Mathematical Statistics* 32: 774-99; September 1961.

Bartholomew, D. J. "Ordered Tests in the Analysis of Variance." *Biometrika* 48: 325-32; December 1961.

Bhapkar, V. P. "Confidence Bounds Connected with ANOVA and MANOVA for Balanced and Partially Balanced Incomplete Block Designs." *Annals of Mathematical Statistics* 31: 741-48; September 1960.

Cox, C. Philip. "The Relation Between Covariance and Individual Curvature Analyses of Experiments with Background Trends." *Biometrics* 18: 12-21; March 1962.

Curnow, R. N., and Sharpe, E. "The Analysis of Covariance as a Means of Reducing Standard Errors in Certain Experiments Involving Sequences of Treatments." *Biometrics* 18: 410-13; September 1962.

Ehrenfeld, S., and Zacks, S. "Randomization and Factorial Experiments." *Annals of Mathematical Statistics* 32: 270-97; March 1961.

Fabian, Václav. "On Multiple Decision Methods for Ranking Population Means." *Annals of Mathematical Statistics* 33: 248-54; March 1962.

Graybill, Franklin A., and Hultquist, Robert A. "Theorems Concerning Eisenhart's Model II." *Annals of Mathematical Statistics* 32: 261-69; March 1961.

Harter, H. Leon. "On the Analysis of Split-Plot Experiments." *Biometrics* 17: 144-49; March 1961.

Hogg, Robert V. "Iterated Tests of the Equality of Several Distributions." *Journal of the American Statistical Association* 57: 579-85; September 1962.

McHugh, Richard B., and Wall, Francis J. "Estimating the Precision of Time Period Effects in Longitudinal Models with Serially Correlated and Heterogeneous Errors." *Biometrics* 18: 520-28; December 1962.

Neyman, Jerzy, and Scott, Elizabeth L. "Correction for Bias Introduced by a Transformation of Variables." *Annals of Mathematical Statistics* 31: 643-55; September 1960.

Pearson, E. S. "Some Thoughts on Statistical Inference." *Annals of Mathematical Statistics* 33: 394-403; June 1962.

Potthoff, Richard F. "Three-Dimensional Incomplete Block Designs for Interaction Models." *Biometrics* 19: 229-63; June 1963.

Roy, S. N., and Cobb, Whitfield. "Mixed Model Variance Analysis with Normal Error and Possibly Non-Normal Other Random Effects, Part I: The Univariate Case." *Annals of Mathematical Statistics* 31: 939-57; December 1960.

Thompson, W. A., Jr. "The Problem of Negative Estimates of Variance Components." *Annals of Mathematical Statistics* 33: 273-89; March 1962.

CHAPTER IV

Regression and Correlation

WILLIAM J. MOONAN and JOHN H. WOLFE

THIS CHAPTER reviews publications relating to regression and correlation theory that have appeared from June 1960 to June 1963. These subjects were last reviewed in the REVIEW by Michael and Hunka in December 1960.

Regression Analysis

Regression Models

A large class of models descriptive of natural and behavioral phenomena can be characterized by an equation of the form $y=F[x,\beta]+e$, where y is the dependent variable of particular interest, $F[x,\beta]$ is a function of some variables x and parameters β, and e is an error component. The model can be particularized in a bewildering number of ways by specifying various kinds of statistical properties of y, x, and e, and forms for $F[x,\beta]$. For example, y and x may be random or fixed variables, and $F[x,\beta]$ may be a linear or nonlinear function (in the usual algebraic sense) of the β. If the $F[x,\beta]$ is β-nonlinear, then statistical literature terms the model *nonlinear*, whereas if $F[x,\beta]$ is x-nonlinear, the model is called *curvilinear*. Semantically, this distinction is not too clear, and it consequently has led to confusion. Nevertheless, it is very important.

The equation $y=\beta_1x_1+\beta_2x_2+e$ is both β-linear and x-linear; whereas $y=\beta_1x_1+\beta_2x_2+\beta_3x_1x_2+e$ is β-linear and x-nonlinear because of the term x_1x_2. The equation $y=\beta_1x_1+(\beta_1)(\beta_2)x_2+e$ is β-nonlinear and x-linear, but $y=\beta_1x_1^2+(\beta_2)(\beta_3)x_2+e$ is β-nonlinear and x-nonlinear. We have made these distinctions at some length because we find this classification system convenient for part of this review. If $y=F[x,\beta]+e$ is looked at as a prediction model, it is usual to assume that y, x, and e are random variables and that $F[x,\beta]$ is β-linear. Also, in this case, if a particular set of values for $x=X$ is given with X and e uncorrelated, and if the elements of e have expected values of zero and if they are independently distributed with constant common variance, then the model is known as one of *multiple regression*.

Particular care needs to be taken to select a proper regression model. Guidance in this matter is difficult to obtain; experience and professional training are the best aids. A much-needed general discussion of model building and of methods of defining, solving, and testing models was presented by Ackoff (1962). Written from the operational analysis point of view, his book contained much useful information about models. Box and Hunter (1962) proposed a technique for building models by a unique experimental procedure that statistically analyzes estimated regression parameters in order to develop, iteratively, better models. A more detailed exposition of the problems associated with estimation and tests of hypotheses of regression and with analysis of variance models was contributed by Graybill (1961). His book is an excellent introduction to linear statistical analysis. Graybill also analyzed the so-called functional model which relates the mathematical variables Y and X. For this case, $F[X,\beta]$ should be X-linear. A very complete exposition of the general theory of regression analysis and of functional model analysis was contributed by Kendall and Stuart (1961). Plackett's (1960) book gave a broad interpretation of regression analysis by defining it as the domain associated with the algebraic theory and with the numerical methods of least squares. Although mathematical, Plackett's work does provide numerous illustrations of computational procedures suitable for desk calculators and computers.

β-Nonlinear Models

If $F[x,\beta]$ is β-nonlinear, estimation problems are generally difficult, mathematically and computationally. A very general β-nonlinear model is the generalized logistic equation defined by Nelder (1961). This model has the exponential curves, the Gompertz equation, the logistic equation, and the Mitscherlich equation as special cases. These forms are associated with regression functions that have an upper and/or lower asymptote. Growth, learning, and many other educational processes have this limiting property. Therefore, these forms, which are very important, have general utility. It is known that x-nonlinear regression analysis is usually not appropriate for data which have single or double asymptotes. Stevens (1951) had previously presented computational procedures using a Taylor expansion for the Mitscherlich equation and its transforms. This computational work was also generalized by Nelder. Turner, Monroe, and Lucas (1961) also gave estimating procedures for a general class of asymptotic models and considered their application to the path analysis of complex causal networks of variables. A modification of the Gauss-Newton iterative procedure for solving simultaneous equations was applied and illustrated for β-nonlinear models by Hartley (1961).

x-Nonlinear Models

The general value of x-nonlinear forms for regression equations is beginning to be recognized in applied educational research. Applications have been prompted in part by the increased facility that computers have provided for making regression analyses. The x-nonlinear forms may involve many kinds and mixtures of functions of a basic set of independent variables. Power functions and crossproducts of variables have proved useful as predictors in research studies. Anderson (1962) investigated the problem of determining what polynomial degree should be included in a regression equation. He provided a procedure for ascertaining whether a higher power than that chosen is appropriate. Ward (1962) has been a resourceful advocate of x-nonlinear analysis. He emphasized that binary variables and such x-nonlinear forms as reciprocals, trigonometric functions, and ratios, which may be useful as potential predictors, are frequently neglected. In addition, he gave a general presentation of regression analysis viewed essentially from the general linear hypothesis standpoint. Bottenberg and Ward (1963) elaborated this concept and gave a refreshingly novel approach to the technique of research problem formulation.

Accretion and Deletion Procedures

From a few original variables, computers permit the generation of a large family of variables as predictors, although the ability to perform the computations does not mean that they should be performed. A researcher should have good theoretical or practical reasons for introducing new or transformed variables into his regression analysis. When this is the case, it is sometimes necessary to engage in some sort of variable selection program. Although many techniques are available for this process, most are variants of the so-called deletion or accretion procedures used by Horst and MacEwan (1960). They applied these methods to problems of multiple absolute and multiple differential prediction and indicated that the deletion procedure is better than the accretion since, at any stage, it utilizes to a greater degree the intercorrelations among the predictor variables.

Larson and Bancroft (1963) contributed an important paper in the study of accretion and deletion procedures. Among other things, they characterized the two procedures in terms of hypothesis testing and showed that under certain conditions, the deletion procedure is superior if the predictions made involve the same predictor scores. Nordø (1960) provided a description of a computer process for a generalized as well as for the simple deletion procedure.

Wolfe, one of the authors of this chapter, has studied polynomial regression and correlations and has found that the order of the accretion of polynomial variables into regression equations depends upon the scaling of the components of the polynomials. He recommends that the components of polynomial variables be mean-deviated. This method prevents polynomials, which contribute little more than their individual component variables, from entering the regression equation prematurely. This action has the additional merit of avoiding some ill-conditioning of the x-variance-covariance matrix. In general, one should be cautious about including polynomials of very high degree in a regression equation. Such terms may add relatively large contributions to prediction estimates, and unless the corresponding regression coefficients are very accurately estimated, large errors can be introduced into the predictions. In such circumstances, cross-validation procedures are especially important.

The problem of shrinkage of the multiple correlation coefficient in cross-validation studies was investigated by Nicholson (1960). He defined an index of the measure of efficiency of prediction equations and gave its distribution function. The confidence interval of this index indicates the degree to which a regression equation derived from an original sample will predict in a crossvalidation sample. He also discussed the index in terms of relative sample sizes and predictors.

A practical discussion of problems associated with the accretion and deletion procedures was furnished by Moore (1962). He pointed out that (a) at any stage of the accretion procedure, an insignificant contributor may appear before a significant variable; (b) the strictly optimum combination of significant predictors may not be discovered; and (c) the application of both procedures to the same data will not necessarily yield the same set of predictors. Moore suggested that the choice of the accretion or deletion method may depend upon the number of predictors desired to be retained in the final regression equation. If K is the number to be retained and V is the number available, he suggested that the accretion procedure should be used if K is very much less than V. Alternatively, if K is only slightly less than V, the deletion procedure should be used. Although this seems to be good advice from a computational standpoint, the method does not obviate the difficulties arising from a complex relationship among the variables. There seems to be no method presently available that minimizes computation and guarantees the optimum selection of predictors. Perhaps there never will be. Although the accretion and deletion methods seem to be reasonably efficient, they are subject to dangers with which sound research and professional judgment will have to cope.

Freund, Vail, and Clunies-Ross (1961), Goldberger and Jochems (1961), and Goldberger (1961) discussed the estimations and significance tests resulting from a two-stage stepwise least-squares accretion procedure. In the first stage of this procedure, a β-linear model is analyzed to determine the usual regression coefficients for the first set of predictors, e.g., x_1. In the second stage, the residuals, $y - \beta_1 x_1 = y_1$, are obtained; and a second regression equation is derived using a second set, x_2, as predictors, β_2 as coefficients, and y_1, the residuals, as criterion. The authors showed that, compared to the usual one-stage analysis of $y = \beta_1 x_1 + \beta_2 x_2 + e$, the estimated parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ and the tests of significance are both biased, and they described the degree and conditions of the biases.

Computer Analysis

Many computer programs are available for performing regression analyses for both β-linear and β-nonlinear models. A surprising number of these programs are redundant, performing the same work, or small variations of it, on the same computers. A good computer regression program will provide for flexible input, an efficient matrix inverter and flexible output to meet a great variety of demands. It is especially important for the programs to allow for the generation of x-nonlinear variables from a set of fundamental x variables. These problems were ably discussed by Healy (1963). Experience has shown, however, that the most general programs are not necessarily the best to use because of the details required to set up particular problems. A good computer library will include a variety of specific regression programs as well as a very general program.

A special program for handling missing x data seems essential. An approach to this problem was suggested by both Buck (1960) and Walsh (1961). Their methods involve the estimation of missing values of x by an independent regression analysis. Since this technique may necessitate a large computational effort, it seems appropriate only if the missing x can be estimated very accurately. Although the resulting x-covariances are unbiased, the x-variances need to be corrected. A simpler procedure is to compute the regression equation from a correlation matrix of all available data and to use the largest sample-size estimate of the means and standard deviations of x. The biases associated with this procedure, which are not known, are related to the factors causing missing x.

Miscellaneous Topics

Within the space allotted for this chapter, it is impossible to cover adequately the many important topics associated with regression theory and practice. The number of references in the current regression literature is very large, and such important topics as correlated errors, weighted regression, and applications to applied science have been neglected. The following sample of references in this section points to the wealth of theoretical and practical information available to interested researchers.

The subject of time series is a fundamental one in the educational process. This fact was pointed out by Brown and Meyer (1961), who characterized a simple learning model in terms of an exponential smoothing process. Their work considers the analysis of a smoothing model for which x may involve polynomials of arbitrary degree. The smoothing process permits forecasts to be made and their errors to be assessed.

Bottenberg and Christal (1961) made an application of regression technique to criterion cluster analysis. They devised a calculation procedure which successively clusters a set of criteria so that, at any stage, there is a minimal loss of predictive efficiency of the clustered criteria by a set of predictors. Such a procedure, which is of great importance for military and academic assignment problems, is a clever application of a deletion process which reduces the number of predictors in sets of variables rather than one at a time.

Rao (1962) developed an important procedure in the theory of regression to apply to selection problems. He reported the theoretical development and the calculation method for constructing a linear function of x which will maximize the correlation with a criterion y so that the same linear functions correlated with other y's will be non-negative. With this methodology, an equation can be obtained which will insure optimum selection with respect to the mean of y_1 under the condition that the means of other criteria will not diminish. For two criteria, the procedure is computationally simple.

Box and Watson (1962) investigated the influence of particular values of predictor variables upon the sensitivity of the F-test in regression analysis to non-normality of the criterion variable. They found that the sensitivity to non-normality in y is determined in great part by the extent of the non-normality of x. Particular significance tests are more sensitive than others. A selected bibliography of about 250 publications (1930-57) relating to regression and correlation theory has been compiled by Deming (1960). Since the source material is world-wide, the list is likely to contain references not previously known to most readers. Although the bibliography is far from being exhaustive, it does contain many important contributions.

Correlation Analysis

Measures of Association

One of the outstanding articles in the past three years was by Goodman and Kruskal (1963), the third paper of a series concerned with measures of association for contingency tables. They presented the sampling theory, maximum-likelihood estimators, and variances for their indexes λ_a, λ_b, λ, γ, and others. The coefficient λ_a, which Goodman and Kruskal credited to Guttman, might be described as the decrease in number of errors in predicting a from b, divided by the number of errors in predicting a without information about b. λ_b is defined similarly. The importance of the measures λ_a and λ_b lies in their asymmetry; since $\lambda_a \neq \lambda_b$, the prediction of a from b may be better than the prediction of b from a. The coefficient λ is a symmetrized combination of λ_a and λ_b, and γ is a measure of association appropriate for ordered categories, closely related to Kendall's tau and equal to Yule's Q in the 2x2 case.

Another type of asymmetry of association arises from so-called implicative relations. For example, success in college requires above-average intelligence, but high intelligence is not sufficient for scholastic achievement. Otherwise stated, prediction is not equally accurate at both ends of the predictor, and there is heteroscedasticity of regression. Maruyama (1962) discussed the problem of discovering and describing implicative relations and made a computational comparison of 20 indexes of association for 15 different 2x2 contingency tables representing different degrees of implication. Among the measures compared were chi square, Yule's Q and Y, Goodman-Kruskal's γ and λ, and 13 new indexes proposed by Maruyama.

Measures of Correlation

Several papers evaluated the performance of various rank correlation coefficients, such as Spearman's correlation of ranks, r_s; Kendall's tau, r_k; and the Fisher-Yates coefficient, r_f (the product-moment correlation of the expected normal deviates corresponding to the ranks). Fieller and Pearson (1961) reported that the Fisher z-transforms of r_s, r_k, and r_f are all nearly normally distributed, but that the Fisher-Yates coefficient has less sample variance and more discriminating power than the other coefficients. Farlie (1960) presented formulas for the mean values and asymptotic relative efficiencies of Spearman's rank correlation and Kendall's tau for a very general class of continuous bivariate distributions.

Several generalizations of the point-biserial correlation coefficient have been proposed. Olkin and Tate (1961) considered the multiple correlation coefficient, the partial correlation coefficients, and the canonical correlations for the case of several dichotomous variables with a multinomial distribution mixed with several continuous variables with conditional multivariate normal distributions. Das Gupta (1960) defined the point-multiserial coefficient for the case of a single continuous variable having different conditional normal distributions for different values of a many-valued discrete variable.

Effects of selection, errors of scaling, and errors of measurement were discussed by Carroll (1961). He showed that when two variables have differently shaped marginal distributions, the maximum possible correlation between them will be less than one; he also illustrated a general method for finding the maximum possible correlation. Carroll noted that errors of scaling and errors of selection can distort the marginal distributions in similar ways, but that the proper adjustments for these two effects may be precisely opposite to each other. Graphs giving the product-moment correlation corrected for restriction in range were presented by Michael and others (1962).

Bibliography

ACKOFF, RUSSELL L. *Scientific Method.* New York: John Wiley & Sons, 1962. 464 pp.

ANDERSON, T. W. "The Choice of the Degree of a Polynomial Regression as a Multiple Decision Problem." *Annals of Mathematical Statistics* 33: 255-65; March 1962.

BOTTENBERG, ROBERT A., and CHRISTAL, RAYMOND E. *An Iterative Technique for Clustering Criteria Which Retains Optimum Predictive Efficiency.* Technical Report WADD-TN-61-30. Lackland Air Force Base, Tex.: Personnel Laboratory, Wright Air Development Division, 1961. 16 pp.

BOTTENBERG, ROBERT A., and WARD, J. H., JR. *Applied Multiple Linear Regression.* Report No. PRL-TDR-63-6. Lackland Air Force Base, Tex.: 6570th Personnel Research Laboratory, Aerospace Medical Division, U.S. Air Force Systems Command, 1963. 139 pp.

BOX, G. E. P., and HUNTER, WILLIAM G. "A Useful Method for Model-Building." *Technometrics* 4: 301-18; August 1962.

BOX, G. E. P., and WATSON, G. S. "Robustness to Non-Normality of Regression Tests." *Biometrika* 49: 93-106; June 1962.

BROWN, ROBERT G., and MEYER, RICHARD F. "The Fundamental Theorem of Exponential Smoothing." *Operations Research* 9: 673-85; September-October 1961.

BUCK, S. F. "A Method of Estimation of Missing Values in Multivariate Data Suitable for Use with an Electronic Computer." *Journal of the Royal Statistical Society* (Series B, Methodological) 22: 302-306; No. 2, 1960.

CARROLL, JOHN B. "The Nature of the Data, or How To Choose a Correlation Coefficient." *Psychometrika* 26: 347-72; December 1961.

DAS GUPTA, S. "Point Biserial Correlation Coefficient and Its Generalization." *Psychometrika* 25: 393-408; December 1960.

DEMING, LOLA S. "Selected Bibliography of Statistical Literature, 1930 to 1957: I. Correlation and Regression Theory." *Journal of Research of the National Bureau of Standards* 64B: 55-68; January-March 1960.

FARLIE, D. J. G. "The Performance of Some Correlation Coefficients for a General Bivariate Distribution." *Biometrika* 47: 307-23; December 1960.

FIELLER, E. C., and PEARSON, E. S. "Tests for Rank Correlation Coefficients: II." *Biometrika* 48: 29-40; June 1961.

FREUND, RUDOLF J.; VAIL, RICHARD W.; and CLUNIES-ROSS, C. W. "Residual Analysis." *Journal of the American Statistical Association* 56: 98-104; March 1961.

GOLDBERGER, ARTHUR S. "Stepwise Least Squares: Residual Analysis and Specification Error." *Journal of the American Statistical Association* 56: 998-1000; December 1961.

GOLDBERGER, ARTHUR S., and JOCHEMS, D. B. "Note on Stepwise Least Squares." *Journal of the American Statistical Association* 56: 105-10; March 1961.

GOODMAN, LEO A., and KRUSKAL, WILLIAM H. "Measures of Association for Cross Classifications: III. Approximate Sampling Theory." *Journal of the American Statistical Association* 58: 310-64; June 1963.

GRAYBILL, FRANKLIN A. *An Introduction to Linear Statistical Models, Volume I.* New York: McGraw-Hill Book Co., 1961. 463 pp.

HARTLEY, H. O. "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares." *Technometrics* 3: 269-80; May 1961.

HEALY, M. J. R. "Programming Multiple Regression." *Computer Journal* 6: 57-61; April 1963.

HORST, PAUL, and MACEWAN, CHARLOTTE. "Predictor Elimination Techniques for Determining Multiple Prediction Batteries." Monograph Supplement 1-V7. *Psychological Reports* 7: 19-50; August 1960.

KENDALL, M. G., and STUART, ALAN. *The Advanced Theory of Statistics: Vol. 2. Inference and Relationship.* Revised edition. New York: Hafner Publishing Co., 1961. 676 pp.

LARSON, HAROLD J., and BANCROFT, T. A. "Sequential Model Building for Prediction in Regression Analysis: I." *Annals of Mathematical Statistics* 34: 462-79; June 1963.

MARUYAMA, MAGOROH. "Conjunctive-Disjunctive-Implicational Model in Discrete Scales." *Journal of Experimental Education* 30: 289-305; March 1962.

MICHAEL, WILLIAM B., and OTHERS. "Abacs for Determination of a Correlation Coefficient Corrected for Restriction of Range." *Psychometrika* 27: 197-202; June 1962.

MOORE, P. G. "Regression as an Analytical Tool." *Applied Statistics* 11: 106-19; June 1962.

NELDER, J. A. "The Fitting of a Generalization of the Logistic Curve." *Biometrics* 17: 89-110; March 1961.

NICHOLSON, GEORGE E., JR. "Prediction in Future Samples." *Contributions to Probability and Statistics.* (Edited by Ingram Olkin and others.) Stanford, Calif.: Stanford University Press, 1960. Chapter 29, pp. 322-30.

NORDØ, JACK. *Significance of Regression Equations Derived from Serially Correlated Data, and a Procedure of Selecting Optimal Predictors.* Technical Note No. 1, ASTIA Document No. AD 245483. Oslo, Norway: Norwegian Meteorological Institute, Geophysics Research Directorate, Air Force Cambridge Research Center, Air Research and Development Center, U.S. Air Force (European Office), 1960. 20 pp.

OLKIN, I., and TATE, R. F. "Multivariate Correlation Models with Mixed Discrete and Continuous Variables." *Annals of Mathematical Statistics* 32: 448-65; June 1961.

PLACKETT, R. L. *Principles of Regression Analysis.* Oxford, England: Clarendon Press, 1960. 174 pp.

RAO, C. RADHAKRISHNA. "Problems of Selection with Restrictions." *Journal of the Royal Statistical Society* (Series B, Methodological) 24: 401-405; No. 2, 1962.

STEVENS, W. L. "Asymptotic Regression." *Biometrics* 7: 247-67; September 1951.

TURNER, MALCOLM E.; MONROE, ROBERT J.; and LUCAS, HENRY L., JR. "Generalized Asymptotic Regression and Non-Linear Path Analysis." *Biometrics* 17: 120-43; March 1961.

WALSH, J. E. "Computer-Feasible Method for Handling Incomplete Data in Regression Analysis." *Journal of the Association for Computing Machinery* 8: 201-11; April 1961.

WARD, JOE H., JR. "Multiple Linear Regression Models." *Computer Applications in the Behavioral Sciences.* (Edited by Harold Borko.) Englewood Cliffs, N.J.: Prentice-Hall, 1962. Chapter 10, pp. 204-37.

CHAPTER V

Factor Analytic Methodology

PHILIP R. MERRIFIELD and NORMAN CLIFF

SINCE 1960, increased attention to epistemological questions has characterized the course of research into factor analytic methods. While interest in the continuing problems of communality, rank, and rotational options has stimulated a number of investigations, the major acceleration emphasized both the design of studies and techniques of factoring in other than the common-factor space underlying matrices of correlations between variables (*R*-technique). The definitive text by Harman (1960) lucidly presented the matrix algebra necessary to understand the common techniques of factor extraction. Examples of several rotational procedures are presented and compared.

In this chapter, the three major sections are concerned with (a) design of studies, (b) decisions within a single study, and (c) selected applications of methods.

Design of Studies

In part because of the increased accessibility of computers, the stream of substantive factor analyses typical of earlier periods threatened to become a torrent as this flow was augmented by the contributions of many whose major purpose seemed to be to "do a factor analysis." On the other hand, the prospects of easier computation than formerly may have contributed to the increased interest in *P*- and *Q*-techniques as well as in the other types of models potentially applicable to psychological problems.

Hypotheses and Interpretations

By combining four basic components—stimuli, responses, occasions, and individuals—Coan (1961) concluded that a total of 24 different possibilities for factor analytic studies was available and suggested ways to analyze these possibilities. Bechtoldt (1961) differentiated between exploratory and confirmatory analyses and pointed out that common criteria such as minimum rank and simple structure may be more appropriate to confirmatory than to exploratory investigations. Bechtoldt's data consisted of scores on 17 tests, each of which was purported to be a univocal measure of one of six abilities, administered to more than 400 subjects.

He divided the group randomly into two samples, each consisting of more than 200 subjects. Factors extracted by centroid, principal axes, and multiple group methods showed little resemblance to each other, whether or not communalities were used. Differences between samples within methods were also marked. However, when guided by prior knowledge of the content of the tests, oblique rotations of all sets of extracted factors produced a remarkable similarity in the rotated solutions. He concluded that a high value of the index of reproducibility of rotated factors was to be expected in a well-designed confirmatory study.

In a later discussion, Bechtoldt (1962) differentiated between the appropriateness of factor analytic and regression models for testing hypotheses. He contended that the inclusion of criterion variables with predictor variables in a factor analysis confuses the relationships. Furthermore, noting the desirability of linear and experimental independence of predictors, one from the other, he suggested extraction by multiple group rather than centroid procedures. He maintained that the former procedure avoids the experimental dependence arising in the latter approach because all the variables are used in the centroid procedure. It must be pointed out that useful results have been obtained when predictors and criteria have been used in the same factor analysis and that experimental independence is as much affected by exclusion of variables, as in multiple group extraction, as it is by their inclusion.

Cattell (1961) discussed similarity in methods of measurement as a source of factors in addition to the usual sources of covariance arising from similarity of traits and from correlation of second-order factors. In a study in which hypotheses were derived from physical similarities among variables, Cattell and Sullivan (1962) demonstrated the degree to which oblique rotations produced factors closely resembling the hypothesized groupings. The reader may find cause for concern in the high correlations of the "error" factor with one "systematic" factor in each of the samples analyzed.

In a long and thorough discussion of the interpretation of rotated factors in oblique solutions, Cattell (1962) presented 6 of the approximately 50 matrices representing possible dimension-variable relationships. He recommended the use of the primary factor pattern as the basis of interpretation, instead of the reference vector structure matrix more commonly reported. He emphasized the need to interpret factors as influences on which the variables in the study are dependent rather than as mere mathematical definitions. For future use, Cattell urged the consideration of a matrix indicating the relationship between a (first-order) factor and a variable, when the effects of higher-order factors are partialed from both the factor and the variable.

It would appear that many varieties of analysis, within the limitations of *R*-technique, are possible; hopefully, one of them will be most appropriate for the investigation of the factor analytic psychologist's multivariate problem.

P-, *Q*-, and *R*-Techniques

Two papers highly critical of *P*-technique (correlations between various measures taken on a single individual over many different occasions) were published. Anderson (1963) questioned the validity of the assumption of independence among error or unique scores. He advised distinguishing in the analysis between independent (controlled or directly manipulated) and dependent (merely observed) variables and between those variables subject to measurement error and those not. He developed the implications of these distinctions for the interpretation of factors and emphasized the value of including time and various functions of time in the analysis. As an alternative to *P*-technique, he proposed a type of analysis which appears to be a combination of time-series analysis (autocorrelation) and multiple correlation and gave references to econometric and statistical literature dealing with similar or related problems. Holtzman (1962) criticized *P*-technique on similar grounds. He too emphasized the lack of independence of observations and the possible presence of serial correlations or lagged serial correlations as well as the resultantly large number of relationships that must be investigated.

In his review of applications of *Q*-methodology, Wittenborn (1961) differentiated between *Q*-technique leading to ipsative measures and *Q*-factor analysis based on correlations between persons over items. He pointed out that ipsative scales are necessary to meet Stephenson's criteria for *Q*-methodology in psychology and noted some of the difficulties in correlating ipsative measures. Sundland (1962) spoke to a crucial point in the correlation of lists of attributes such as those commonly used in obtaining *Q*-sorts, where synonyms or closely associated labels are included in the same list. The synonyms, considered as replications over which the covariance is an average, are not experimentally independent. Sundland alleged that the computed standard errors of the correlations, or of the trait estimates, are smaller than their "true" values. This notion would appear to warrant further investigation, especially as it bears on the reliability of trait scores derived from *Q*-analysis.

To investigate the utility of carrying out an ipsative process upon normative scores in lieu of obtaining experimentally ipsative scores, Wright (1961) administered the *Edwards Personal Preference Schedule* in rating scale form and derived normative and normative-ipsative measures for comparison with the usual interactive scores. Interscale correlations were factored separately for the three measures; factors from normative and normative-ipsative scores resembled each other more closely than either the normative or normative-ipsative factors resembled the factors from the experimentally ipsative scores. It should be noted that all three analyses were carried out using *R*-technique. It would be interesting to compare interperson correlations based on the three kinds of measures.

In his exploration of the relationships between *Q*- and *R*-techniques, Broverman (1961) pointed out that the results depend on whether rows or columns are standardized in the raw-score matrix. However, in a later article, Broverman (1962) noted the lack of identity of the factors and inferred a lack of any relationship—a misstep noted by MacAndrew and Forgy (1963) and subsequently clarified by Broverman (1963). Addressing the problem of a suitable correlation coefficient for Q-analysis, Guilford (1963) compared several possibilities and recommended a new statistic appropriate for use with item responses or other dichotomous raw data. The statistic is the Pearson-r for numbers that are standard scores of persons relative to item means and item standard deviations. Guilford noted that further investigation of the distribution properties of the new statistic was needed.

Other Techniques and Models

Horst (1961) reviewed the problem of determining the linear combinations of variables from one set which correlate maximally with variables in another set and extended the procedures to more than two sets of variables. Several implications for factoring were drawn. Harris (1962) pointed out some relationships between Rao's canonical factor analysis and Guttman's image analysis. In a somewhat similar article, Jöreskog (1962) advocated the use of residuals as determiners of the rank of the matrix and presented procedures for estimating factor scores and factor loadings. His residuals appear to be related nonlinearly to Guttman's anti-images (which are defined as unity less squared multiple correlations). Guttman (1960) presented five matrices based on the images of the variables and concluded that parsimony, in the sense of the least number of parameters needed per observed variable, does not require the reduction of the R-matrix to its minimum rank.

As a method for using factor analysis to study preferences of individuals for objects, Coombs and Kao (1960) discussed the distance between person and object as obtained in the method of multidimensional unfolding as a suitable measure for factor analysis. They conjectured the appearance of an extra "social utility" dimension corresponding to the distance of each person or object point from the centroid of all such points in the same space. Their conjectures were partially confirmed by two hypothetical examples. Cliff and Ross (1962) showed that the conjecture applies exactly when squared distances are used.

Four papers dealt with Guttman's simplex and circumplex models. Kaiser (1962b) not only showed that paired-comparison procedures may be applied to correlations between variables to order them on an underlying continuum but also provided an index of goodness of fit of the matrix to a simplex. Jones (1962) discussed several methods of analyzing correlation matrices having simplicial form. A different kind of process was implied in each method. Although the intertrial correlations from learning-curve data are shown to exhibit a sort of simplex, Jones's procedures contained no guards against Type II errors. Moreover, he seemed to mistake the apparent simplicity (skill) of practiced motor performance for simplicity in the structure of individual differences between learners. Humphreys (1960) presented data showing that correlations between scores on the same variable measured at different times in the individual's life or at different stages of training have the form of the simplex. However, he pointed out that matrices that simply look like simplexes can sometimes be generated by two common factors of the usual kind. Closer inspection is necessary in order to decide whether a simplex exists. Methods for analyzing a given set of correlations into simplex components are included. Gibson (1963) provided an analysis of the circumplex, in which the correlations, rather than decreasing steadily from the main diagonal as in a simplex, first decrease and then increase toward the off-diagonal corners. Guttman proposed a pattern of factor loadings leading to such a correlation matrix. Gibson furnished a method for recovering a pattern of factor loadings from a correlation matrix, if one may presume that is indeed a circumplex.

Gibson (1962a) discussed the problem of assigning new data to previously found latent classes. Using a procedure similar to the Dwyer extension in factor analysis, he provided a chi-square test for the goodness of fit of the appropriate joint proportions. In an earlier paper, Gibson (1960a) addressed the possibility that latent classes may be interpreted as nonlinear relationships between factors—at least for certain kinds of data.

In a direct attack on the problems of nonlinear factor analysis, McDonald (1962) presented methods for finding out whether a given set of obtained factors actually include a smaller number of independent continuums—the remaining factors being composites of first- and second-order functions of the independent factors.

Additional References: Battig (1962); Bock (1960); Du Bois (1962); Foa (1961); Horst (1962); Marks, Michael, and Kaiser (1961a, b); Miller, Eyman, and Dingman (1961); Rust, Jones, and Kaiser (1962).

Decisions Within a Single Study

Having decided upon the hypotheses he wishes to investigate, the tyro factor analyst typically proceeds to collect his data and thus bypasses an important aspect—the development of suitable tests. Unfortunately, no articles dealing primarily with the psychological aspects of tests to be factor analyzed could be found. Presuppositions regarding the factorial complexity of tests as well as hypotheses regarding their factorial validity may well influence the investigator both in his choice of communality estimates—a choice which thereby influences the rank of his reduced matrix—and in his decisions regarding not only the obliqueness of the rotated factors but also the most acceptable simple structure in the rotated factor space.

Communality and Rank

It is often surmised that the factoring of different coefficients of correlation, although based on the same raw data, will result in different values of communalities and perhaps in differences in the rank of the reduced matrix. The factorial complexity of a test is related both to its communality and to the rank of the matrix in which it is analyzed.

In his discussion of the effects of disparate marginal distributions, Carroll (1961) pointed out that wide variations may induce "difficulty" factors which may or may not be orthogonal to other factors of greater interest. If the investigator is willing to assume that the variables involved have a normal distribution in the population from which the samples at hand were drawn, then, Carroll concluded, the tetrachoric coefficient is least biased. Henrysson (1962) developed the relationship that loadings on the first centroid of a matrix of interitem phi (or tetrachoric) coefficients are biserial coefficients of correlation of item with factor (or test). He suggested the use of factor analysis of items in the development of homogeneous tests.

Now one may question whether person scores on homogeneous tests, which are developed by the application of Henrysson's findings, would have a normal distribution. There is scattered evidence that the scores of persons on highly homogeneous tests may have a skewed rather than a normal distribution. Skewness appears to be accentuated if the same persons' responses to items were analyzed in order to group items into homogeneous tests. The next consideration is whether normality would exist in the population to permit the use of the tetrachoric coefficient even if the sample distributions were markedly skewed, or whether the more conservative assumption resulting in the use of phi coefficients is preferable, as in the study reported by Merrifield, Guilford, and Gershon (1963).

Although the use of unity as a diagonal entry is required if one wishes to obtain uncorrelated factor scores, as pointed out by Overall (1962), spurious factors may be engendered by using unity as an initial diagonal entry if the true values of communality are greatly less than unity. Apparently the interrelations of communality, rank, coefficient of correlation, and homogeneity of tests require further study.

Gibson (1961) suggested the partitioning of the correlation matrix into asymmetric submatrices and derived communality estimates from the intersubmatrix correlations and thus extended Tucker's earlier procedure. However, because of the danger that one or more of the submatrices may be singular—especially in small empirical matrices—the procedure should not be employed blindly.

Comrey (1962) reported his successful use of a new method of extraction. He described an estimated vector, the inner product of which with a row of the correlation matrix is the numerator of a ratio. The denominator of this ratio is the sum of squares of factor estimates for all other variables in the matrix. He used, as an initial estimate of factor vector values, the highest correlation in the column. The computation required for the estimates to converge is more than that needed for centroid analysis but less than that required for principal components. If the factor vector values are considered as beta weights, the numerator of Comrey's ratio becomes the squared multiple correlation. The interesting combination of this lower-bound estimate of communality and the long-used "highest in column" estimate may have the merits of both estimates.

Kaiser (1961) provided an extension of Guttman's lower bound concerning the number of common factors in a correlation matrix to the case in which communalities may lie in the closed interval from zero to unity.

In a penetrating article, Maxwell (1961) reviewed the centroid, the principal-component, and the maximum-likelihood methods of extracting factors and presented comparative data on their application. Tests of significance for the number of factors were reviewed; most of these tests incorporate the chi-square statistic. Maxwell questioned the efficacy of rotation of factors—primarily on the grounds that no adequate sampling theory underlies the notions associated with simple structure. Further, he decried not only the lack of concern with such relevant variables as maturation, sex, and environment but also the inadequacies of samplings of persons.

Rotational Options

In contrast to the welcome accorded the *deus ex machina* of analytic-rotational methods via computer several years ago, the current climate encourages the growth of procedures for using prior information in reaching rotational solutions. Fortunately, this utilization, which can be programed, results in a combination of mechanical and informational approaches. (These terms do seem more indicative of current thinking than do the older terms, *analytic* and *subjective*.) The dangers of hypothesizing after the fact persist—a circumstance which is a natural, not a scientific, attribute.

Lawley (1960) presented a procedure for computing rotated factor loadings directly from the correlation matrix. He used the device of specifying in an appropriate matrix where the zeroes were to occur. Concerning the invariance aspect of rotation, Cattell and Baggaley (1960) developed an index for matching factors in two or more studies. The index constitutes an application of the hypergeometric (or Fisher's exact) probability function, and its effective use depends in large part on the choice of a significance level for factor loadings. To its credit, it avoids the anachronism of computing Pearson-*r* coefficients on loadings intentionally rotated to a highly skewed and truncated distribution. As mentioned previously, Horst (1961) started from the case of two sets of factors (for example, factor matrices for the same variables from two studies) and derived a solution which is the same as Hotelling's "most-predictable criterion." If there are more than two sets, however, the methods must be revised somewhat, but remain essentially similar in that they are derived from a least-squares criterion. Ahmavaara (1963) took a different approach, mainly because he required that each of the transformations relating one set of data to another belong to the same group, in the strict mathematical sense.

Cattell and Gorsuch (1963) added random components to an established simple-structure matrix and maintained the communalities and relative sizes of the factors. Then they applied an arbitrary orthogonal rotation to both factor matrices, and finally they rotated each matrix to simple structure. They concluded that substantially better simple structure obtained for the "real" than for the "randomized" data. Cattell and Dickman (1962) used artificial data to investigate the results of orthogonal and oblique rotations of factors derived from a number of measurements on a number of balls. Factor analysis of the data yielded the expected factors. Oblique rotation resulted in what would be judged to be a good simple-structure matrix in which the number of zero loadings was significant by Bargman's test. Although neither qualitatively satisfactory nor statistically significant in terms of hyperplane counts because of the presence of many intermediate-sized loadings, the orthogonal solution (varimax) resulted in factors yielding the same interpretations as those in the oblique solution.

In two papers, Gibson (1960b, 1962b) addressed the problem of obtaining orthogonal from oblique transformations. In his first contribution, he gave a geometric comparison of orthogonal and oblique factors and of their interrelationship; in the second, he showed how to find an orthogonal transformation having a least-squares fit to a given oblique transformation. Presumably, the orthogonal matrix would be applied to the unrotated factor matrix to obtain an orthogonal rotated solution. Gocka (1962) presented the results of an empirical investigation on the optimal weights for Thurstone's analytic rotation method.

In the area of rotational criteria, Saunders (1961) extended quartimax notions to the oblique case and called the resulting procedure "oblimax." In their development of a program to maximize the number of zero entries in each hyperplane, Cattell and Muerle (1960) noted that the oblique factors meeting their "maxplane" criterion required visual improvement to meet all psychological criteria. Eyman, Dingman, and Meyers (1962) concluded that biquartimin produced more meaningful results than oblimax for the oblique case and that varimax produced more readily interpretable results than quartimax for the orthogonal case. Although they could not obtain a satisfactory orthogonal solution for their data by the Zimmerman graphic procedure, they did achieve a satisfactory Thurstone oblique solution for comparison with the mechanical procedures. Their criteria were simple structure and psychological meaning. Marks, Michael, and Kaiser (1960) applied the varimax criterion to a matrix of aptitude variables for which a satisfactory informational solution had been achieved.

As a result, the number of factors having significant loadings was greatly reduced, and the factors themselves represented composites of those obtained in the informational solution. Merrifield, Guilford, and Gershon (1963) reported a similar finding with less distortion than in the previously mentioned study. The varimax procedure, which may have been required to rotate too many factors, did result in larger groups of tests and in more factors having but one significant loading than did the graphical solution. In the use of varimax, it seems important that each factor be represented by three or more tests and that the number of centroid factors or principal components to be rotated be specified correctly.

In their extensive discussion of the rotational problems arising from sampling of variables, Guilford and Zimmerman (1963) concluded that the investigator has a responsibility to use information regarding the previous performance of the measures in the battery, the factorial complexity where it is known or suspected in a test, and the degree to which a factor is represented in a battery. Informational solutions are preferred to mechanical; and orthogonal, to oblique.

Additional References: Burroughs and Miller (1961); Gibson (1961); Howard and Cartwright (1962); Howard and Gordon (1963); Jarrett (1961); Lumsden (1961); McQuitty (1961a, b); Saunders (1962). Because of limitations of space, computer programs for factor analysis cannot be discussed in detail. Selected examples may be found in Dixon (1962a, b); Hurley and Cattell (1962); Saunders (1960); Vandenberg and Lamphiear (1960); and Wexler (1960).

Selected Applications of Methods

The citations in this section were selected for methodological interest. Kaiser (1962a) indicated how component scores might be obtained from principal components; rotated factors were not considered in his discussion. Kaiser and Dickman (1962) developed score matrices and correlation matrices by sampling from a known population of factor scores and correlations; this type of effort should be helpful in studying the effects of sampling on factorial invariance.

Dingman and McIntyre (1962) showed that the loadings of raters on factors derived from interrater correlations could be used to indicate when different raters were using different criteria and illustrated how this information might be applied to improve selection of employees.

In comparing factor analytic procedures with profile analysis, Nunnally (1962) demonstrated that the factor analysis of raw-score distance matrices would lead to useful profile information. He noted that factor analysis as a technique is not restricted to the analysis of correlation matrices but that it applies to any Gramian matrix. Nunnally also recommended the application of factor analysis for obtaining clusters of homogeneous profiles. Horn (1961) suggested the use of tests developed in the context of discriminant-function analysis to ascertain the significance of various profile statistics. Cliff (1962) and Weitzman (1963) went even further afield. Cliff showed that factor analysis applied to matrices of raw data may be used in testing the appropriateness of a variety of explicit mathematical theories. Weitzman presented methods for analyzing trials-by-subjects learning data into idealized learning curves. One advantage of this approach is that the total learning curves, rather than single characteristics of them, can be compared for different treatment groups. He provided procedures for estimating different components of the learning curves as well as statistical tests for comparing different learning curves.

Additional References: Haertzen (1963); Harris (1962); Sawrey, Keller, and Conger (1960).

Summary

Contributions to factor analytic methodology since 1960 were reviewed under three headings: design of studies, decisions within a single study, and selected applications. There was increased attention to the specification of the purposes of factor analysis and to the rules for the interpretation of obtained factors. Several investigators explored the persistent problem areas involving communality, coefficients of correlation, and rank. If the present trend continues, new rotational methods should evolve as combinations of mechanical and informational approaches—methods that not only add computer capabilities to those of the factor analyst but also stimulate him to develop inferential procedures consistent with both sampling variability and existing knowledge.

Bibliography

AHMAVAARA, Y. *On the Mathematical Theory of Transformation Analysis.* Report No. 1. Helsinki, Finland: Suomalaisen Kirjallisuuden Kirjapairo Oy Helsinki, 1963. 11 pp.

ANDERSON, T. W. "The Use of Factor Analysis in the Statistical Analysis of Multiple Time Series." *Psychometrika* 28: 1-25; March 1963.

BECHTOLDT, HAROLD P. "An Empirical Study of the Factor Analysis Stability Hypothesis." *Psychometrika* 26: 405-32; December 1961.

BECHTOLDT, HAROLD P. "Factor Analysis and the Investigation of Hypotheses." Monograph Supplement 2-V14. *Perceptual and Motor Skills* 14: 319-42; April 1962.

BROVERMAN, DONALD M. "Effects of Score Transformations on Q and R Factor Analysis Techniques." *Psychological Review* 68: 68-80; January 1961.

BROVERMAN, DONALD M. "Normative and Ipsative Measurement in Psychology." *Psychological Review* 69: 295-305; July 1962.

BROVERMAN, DONALD M. "Comments on the Note by MacAndrew and Forgy." *Psychological Review* 70: 119-20; January 1963.

CARROLL, JOHN B. "The Nature of the Data, or How To Choose a Correlation Coefficient." *Psychometrika* 26: 347-72; December 1961.

CATTELL, RAYMOND B. "Theory of Situational, Instrument, Second Order, and Refraction Factors in Personality Structure Research." *Psychological Bulletin* 58: 160-74; March 1961.

Cattell, Raymond B. "The Basis of Recognition and Interpretation of Factors." *Educational and Psychological Measurement* 22: 667-97; Winter 1962.

Cattell, Raymond B., and Baggaley, Andrew R. "The Salient Variable Similarity Index for Factor Matching." *British Journal of Statistical Psychology* 13: 33-46; May 1960.

Cattell, Raymond B., and Dickman, Kern. "A Dynamic Model of Physical Influences Demonstrating the Necessity of Oblique Simple Structure." *Psychological Bulletin* 59: 389-400; September 1962.

Cattell, Raymond B., and Gorsuch, Richard L. "The Uniqueness and Significance of Simple Structure Demonstrated by Contrasting Organic 'Natural Structure' and 'Random Structure' Data." *Psychometrika* 28: 55-67; March 1963.

Cattell, Raymond B., and Muerle, John L. "The 'Maxplane' Program for Factor Rotation to Oblique Simple Structure." *Educational and Psychological Measurement* 20: 569-89; Fall 1960.

Cattell, Raymond B., and Sullivan, William. "The Scientific Nature of Factors: A Demonstration by Cups of Coffee." *Behavioral Science* 7: 184-93; April 1962.

Cliff, Norman. "Analytic Rotation to a Functional Relationship." *Psychometrika* 27: 283-95; September 1962.

Cliff, Norman, and Ross, John. "A Generalization of the Interpoint Distance Model." (Abstract) *American Psychologist* 17: 365-66; June 1962.

Coan, Richard W. "Basic Forms of Covariation and Concomitance Designs." *Psychological Bulletin* 58: 317-24; July 1961.

Comrey, Andrew L. "The Minimum Residual Method of Factor Analysis." *Psychological Reports* 11: 15-18; August 1962.

Coombs, Clyde H., and Kao, Richard C. "On a Connection Between Factor Analysis and Multidimensional Unfolding." *Psychometrika* 25: 219-31; September 1960.

Dingman, Harvey F., and McIntyre, R. B. "Determining Weights for Composite Ratings by Factor Analysis." *Psychological Reports* 10: 475-80; April 1962.

Eyman, R. K.; Dingman, H. F.; and Meyers, C. E. "Comparison of Some Computer Techniques for Factor Analytic Rotation." *Educational and Psychological Measurement* 22: 201-14; Spring 1962.

Gibson, W. A. "Non Linear Factors in Two Dimensions." *Psychometrika* 25: 381-92; December 1960. (a)

Gibson, W. A. "Orthogonal from Oblique Transformations." *Educational and Psychological Measurement* 20: 713-21; Winter 1960. (b)

Gibson, W. A. "An Asymmetric Approach to Multiple-Factor Analysis." *British Journal of Statistical Psychology* 14: 97-107; November 1961.

Gibson, W. A. "Extending Latent Class Solutions to Other Variables." *Psychometrika* 27: 73-81; March 1962. (a)

Gibson, W. A. "On the Least-Squares Orthogonalization of an Oblique Transformation." *Psychometrika* 27: 193-95; June 1962. (b)

Gibson, W. A. "Factoring the Circumplex." *Psychometrika* 28: 87-92; March 1963.

Gocka, Edward F. "An Algebraic Computation of Weights for Thurstone's Analytic Rotation Method." *Psychological Reports* 11: 507-14; October 1962.

Guilford, J. P. "Preparation of Item Scores for the Correlations Between Persons in a *Q* Factor Analysis." *Educational and Psychological Measurement* 23: 13-22; Spring 1963.

Guilford, J. P., and Zimmerman, Wayne S. "Some Variable-Sampling Problems in the Rotation of Axes in Factor Analysis." *Psychological Bulletin* 60: 289-301; May 1963.

Guttman, Louis. "The Matrices of Linear Least-Squares Image Analysis." *British Journal of Statistical Psychology* 13: 109-18; November 1960.

Harman, Harry H. *Modern Factor Analysis.* Chicago: University of Chicago Press, 1960. 469 pp.

Harris, Chester W. "Some Rao-Guttman Relationships." *Psychometrika* 27: 247-63; September 1962.

Henrysson, Sten. "The Relation Between Factor Loadings and Biserial Correlations in Item Analysis." *Psychometrika* 27: 419-24; December 1962.

Holtzman, Wayne H. "Methodological Issues in *P* Technique." *Psychological Bulletin* 59: 248-56; May 1962.

Horn, J. L. "Significance Tests for Use with r_p and Related Profile Statistics." *Educational and Psychological Measurement* 21: 363-70; Summer 1961.

HORST, PAUL. "Relations Among *m* Sets of Measures." *Psychometrika* 26: 129-49; June 1961.

HUMPHREYS, LLOYD G. "Investigations of the Simplex." *Psychometrika* 25: 313-23; December 1960.

JONES, MARSHALL B. "Practice as a Process of Simplification." *Psychological Review* 69: 274-94; July 1962.

JÖRESKOG, K. G. "On the Statistical Treatment of Residuals in Factor Analysis." *Psychometrika* 27: 335-54; December 1962.

KAISER, HENRY F. "A Note on Guttman's Lower Bound for the Number of Common Factors." *British Journal of Statistical Psychology* 14: 1-2; May 1961.

KAISER, HENRY F. "Formulas for Component Scores." *Psychometrika* 27: 83-87; March 1962. (a)

KAISER, HENRY F. "Scaling a Simplex." *Psychometrika* 27: 155-62; June 1962. (b)

KAISER, HENRY F., and DICKMAN, KERN. "Sample and Population Score Matrices and Sample Correlation Matrices from an Arbitrary Population Correlation Matrix." *Psychometrika* 27: 179-82; June 1962.

LAWLEY, D. N. "Approximate Methods in Factor Analysis." *British Journal of Statistical Psychology* 13: 11-17; May 1960.

MACANDREW, CRAIG, and FORGY, EDWARD. "A Note on the Effects of Score Transformations in *Q* and *R* Factor Analysis Techniques." *Psychological Review* 70: 116-18; January 1963.

MCDONALD, RODERICK P. "A General Approach to Non Linear Factor Analysis." *Psychometrika* 27: 397-415; December 1962.

MARKS, ALVIN; MICHAEL, WILLIAM B.; and KAISER, HENRY F. "Comparison of Manual and Analytic Techniques of Rotation in a Factor Analysis of Aptitude Test Variables." *Psychological Reports* 7: 519-22; August 1960.

MAXWELL, A. E. "Recent Trends in Factor Analysis." *Journal of the Royal Statistical Society* (Series A, General) 124: 49-59; No. 1, 1961.

MERRIFIELD, P. R.; GUILFORD, J. P.; and GERSHON, A. *Differentiation of Divergent-Production Abilities at the Sixth-Grade Level.* Reports from Psychological Laboratory, No. 27. Los Angeles: Department of Psychology, University of Southern California, 1963. 16 pp.

NUNNALLY, JUM. "The Analysis of Profile Data." *Psychological Bulletin* 59: 311-19; July 1962.

OVERALL, JOHN E. "Orthogonal Factors and Uncorrelated Factor Scores." *Psychological Reports* 10: 651-62; June 1962.

SAUNDERS, D. R. "The Rationale for an 'Oblimax' Method of Transformation in Factor Analysis." *Psychometrika* 26: 317-24; September 1961.

SUNDLAND, DONALD M. "The Construction of *Q* Sorts: A Criticism." *Psychological Review* 69: 62-64; January 1962.

WEITZMAN, R. A. "A Factor Analytic Method for Investigating Differences Between Groups of Individual Learning Curves." *Psychometrika* 23: 69-80; March 1963.

WITTENBORN, J. R. "Contributions and Current Status of *Q* Methodology." *Psychological Bulletin* 58: 132-42; March 1961.

WRIGHT, CALVIN E. "A Factor Dimension Comparison of Normative and Ipsative Measurements." *Educational and Psychological Measurement* 21: 433-44; Summer 1961.

Additional References

BATTIG, WILLIAM F. "Parsimony in Psychology." Monograph Supplement 3-V11. *Psychological Reports* 11: 555-72; October 1962.

BOCK, R. DARRELL. "Components of Variance Analysis as a Structural and Discriminal Analysis for Psychological Tests." *British Journal of Statistical Psychology* 13: 151-63; November 1960.

BURROUGHS, G. E. R., and MILLER, H. W. L. "The Rotation of Principal Components." *British Journal of Statistical Psychology* 14: 35-49; May 1961.

DIXON, W. J. "BIMED 17. Factor Analysis." *Behavioral Science* 7: 266; April 1962. (a)

DIXON, W. J. "BIMED 19. Varimax Rotation." *Behavioral Science* 7: 267; April 1962. (b)

DU BOIS, PHILIP H. "On Relationships Between Numbers and Behavior." *Psychometrika* 27: 323-33; December 1962.

FOA, URIEL G. "Convergences in the Analysis of the Structure of Interpersonal Behavior." *Psychological Review* 68: 341-53; September 1961.

GIBSON, W. A. "A Retraction on Inter-Battery Factor Analysis." *Psychometrika* 26: 451-52; December 1961.

HAERTZEN, CHARLES A. "Method for Direct Determination of Inverted Factor Loadings." *Psychological Reports* 12: 399-402; April 1963.

HARRIS, CHESTER W. "Some Problems in the Description of Change." *Educational and Psychological Measurement* 22: 303-19; Summer 1962.

HORST, PAUL. "Matrix Reduction and Approximation to Principal Axes." *Psychometrika* 27: 169-78; June 1962.

HOWARD, KENNETH I., and CARTWRIGHT, DESMOND S. "An Empirical Note on the Communality Problem in Factor Analysis." *Psychological Reports* 10: 797-98; June 1962.

HOWARD, KENNETH I., and GORDON, ROBERT A. "Empirical Note on the 'Number of Factors' Problem in Factor Analysis." *Psychological Reports* 12: 247-50; February 1963.

HURLEY, JOHN R., and CATTELL, RAYMOND B. "The Procrustes Program: Producing Direct Rotation To Test a Hypothesized Factor Structure." *Behavioral Science* 7: 258-62; April 1962.

JARRETT, R. F. "A Note on Rotation and Psychological Space." *British Journal of Statistical Psychology* 14: 109-15; November 1961.

LUMSDEN, JAMES. "The Construction of Unidimensional Tests." *Psychological Bulletin* 58: 122-31; March 1961.

MCQUITTY, LOUIS L. "Elementary Factor Analysis." *Psychological Reports* 9: 71-78; August 1961. (a)

MCQUITTY, LOUIS L. "Substantive Vector Analysis: An Alternative or an Adjunct to Simple Structure." *Psychological Reports* 9: 347-58; October 1961. (b)

MARKS, ALVIN; MICHAEL, WILLIAM B.; and KAISER, HENRY F. "Comparison of Factored Dimensions of Peer and Superior Ratings on a Situational-Type Rating Scale of Intellectual Aptitude." *Psychological Reports* 8: 259-62; April 1961. (a)

MARKS, ALVIN; MICHAEL, WILLIAM B.; and KAISER, HENRY F. "Dimensions of Creativity and Temperament in Officer Evaluation." *Psychological Reports* 9: 635-38; December 1961. (b)

MILLER, CURTIS R.; EYMAN, RICHARD K.; and DINGMAN, HARVEY F. "Factor Analysis, Latent Structure Analysis and Mental Typology." *British Journal of Statistical Psychology* 14: 29-34; May 1961.

RUST, VELMA I.; JONES, R. STEWART; and KAISER, HENRY F. "A Factor-Analytic Study of Critical Thinking." *Journal of Educational Research* 55: 253-59; March 1962.

SAUNDERS, D. R. "A Basic IBM 650 Program To Find Best-Fitting Orthogonal Factors for a Given Hypothesis." *Behavioral Science* 5: 272; July 1960.

SAUNDERS, DAVID R. "Integrating the Implementation of Quartimax, Varimax, Oblimax, and Related Rotational Methods." *Psychological Reports* 10: 241-42; February 1962.

SAWREY, WILLIAM L.; KELLER, LEO; and CONGER, JOHN J. "An Objective Method of Grouping Profiles by Distance Functions and Its Relation to Factor Analysis." *Educational and Psychological Measurement* 20: 651-73; Winter 1960.

VANDENBERG, STEVEN G., and LAMPHIEAR, DONALD. "Varimax and Quartimax Rotations on the IBM 650." *Behavioral Science* 5: 99; January 1960.

WEXLER, JONATHAN D. "MAPS, A System of Interrelated Programs for Factor Analysis on the IBM 704." *Behavioral Science* 5: 272; July 1960.

CHAPTER VI

Design of Sample Surveys in Education

FRANCIS G. CORNELL and EUGENE P. MC LOONE *

ALTHOUGH applications of modern sampling methods are growing in education, fields other than education remain in the forefront both in methodology and in applications. Few sample surveys in education exhibit the sophistication found in studies carried out in other fields. Application of sample surveys to educational situations and areas has been largely by persons from disciplines other than education. Consequently, much of the technical literature on the subject deals with matter outside the field of education.

General References

The importance of sample surveys for educational research and the developments in this field since the detailed presentation by Sitgreaves and Solomon in the December 1957 REVIEW justify a separate chapter in this issue. This chapter supplements a section of the one by Michael and Hunka in the December 1960 REVIEW. In the *Encyclopedia of Educational Research*, Cornell (1960) described the major methods of sampling.

Nontechnical Sources

Sample survey methods apply to the problem of enumeration—in other words, to the determination of how many or how much. There are now excellent sources of relatively nontechnical information on sampling theory and methods, which should be useful to researchers in education who are not themselves mathematical statisticians. Necessary steps are clearly defined by Slonim (1960) as follows:

1. Determine, as precisely as possible, the population or universe to be studied.
2. Set up a sample frame, that is, a list or file of the units in the statistical universe.
3. Designate the information desired from the survey.
4. Determine the desired degree of precision in terms of permissible error (tolerance) and acceptable risk (confidence).

---

* Eugene P. McLoone is specialist, Economics of School Finance, Office of Education, U.S. Department of Health, Education, and Welfare. The interpretations and opinions expressed are those of the author and do not necessarily reflect those of the Office of Education or the Department of Health, Education, and Welfare.

5. On the basis of the desired degree of precision, select the appropriate sample design.
6. Design the reporting form.
7. Conduct a small-scale pretest of the reporting instrument.
8. The report form may be reconstructed on the basis of pretest results.
9. Undertake the survey as planned.

While essential to a sample survey, these steps do not have to be followed precisely in this order. Obviously, there can be overlapping, and one step may precede another.

The Phi Delta Kappa Symposium on Research Design and Analysis outlined sample design methods and explored difficulties in sample surveys in education. The chapter by Kish (1961) is an excellent introduction to sample design theory. The three difficulties of the reaction of the public, the training of researchers, and the technology in education were cited in a chapter by Cornell (1961). Acceptance of sample surveys by the public and by research workers in education should develop from the popular treatment of sampling by Slonim.

Development of Statistical Theory

In a general sense, sample surveys are a special application of the central problem of theory and methods of statistical inference, namely, regulating the degree of uncertainty. The idea of randomization, originally introduced by Fisher, by which the statistician introduces an element of chance into the selection of observations and thus creates a random experiment is shown by Dalenius (1962) to be a link between several branches of design for collecting and analyzing such observations as experiment, acceptance sampling, Monte Carlo sampling, and sample survey. In reviewing developments since 1950 in survey theory and methods, Dalenius pointed to similarities in developments prior to that time. Though efforts continue in the development of sampling variance models, increasing attention is being given to survey theory and to methods that take into account nonsampling errors, what have come to be called mixed error models. This attention to nonsampling errors is a development of great importance in educational surveys, since more often than not a mixed error model is needed.

To attain technical proficiency in survey sampling, the educational statistician should begin with the standard works by Cochran; Deming; Hansen, Hurwitz, and Madow; and Sukhatme and Yates, to which reference was made in earlier issues of the REVIEW. With these basic texts and the Dalenius paper as background, access to such journals as *Biometrika*, the *Annals of Mathematical Statistics*, and the *Journal of the American Statistical Association* should provide leads for keeping up in the field. Short of this effort, the educational researcher would best enlist the services of a mathematical statistician experienced both in the theory and in the practice of sample survey methods.

Among the books that have appeared in the period covered by this REVIEW is a revision of Cochran's (1963) excellent text and a new book on sample design in business research by Deming (1960). The applications of the procedures described by Deming to the managing and operating of schools are legion.

Sample Surveys in Education

The use of sample surveys in educational research is growing. Mathematical statisticians have been employed by the U.S. Office of Education for assistance in sample design since 1960.

In conjunction and in cooperation with various groups, the National Science Foundation has conducted a number of studies of interest to educators. With the Bureau of Social Science Research, it carried out a study of the 1958 college graduating class and a follow-up investigation of the 1958 graduates enrolled in graduate school. In association with the National Opinion Research Center (NORC) under a cooperative research project in the U.S. Office of Education, it completed a study of career choices and decisions of 1961 graduates (Davis, 1963). Presently with NORC, the Foundation is studying the cooperation of graduate education by various fields. In cooperation with the National Science Teachers Association (NSTA), the National Science Foundation (1963) completed studies of the salaries, education, and class loads of science and mathematics teachers in the secondary schools. The employment of mathematicians in industry and their education and training were also subjects studied by the National Science Foundation (1962). A sample of 70,000 persons, who were selected by occupation from the 1960 Census of Population, is being conducted by the Foundation to examine the training and education of persons in various occupations.

As described in a recent report by Anderson and Hoofnagle (1960), the U.S. Department of Agriculture is continuing its study of school food services on a sample basis. Similarly, the U.S. Department of Commerce, Bureau of the Census (1958, 1959, 1960a,b, 1962a,b,c, 1963a,b,c) continued its studies of school enrollment, school employment, school finance, and illiteracy.

Five studies in which sample survey methods were used were issued by the U.S. Office of Education. Approximately 25 additional studies, in which various sample designs are being employed, are in progress. Sample surveys published by the Office dealt with teacher turnover, the status of beginning teachers, programs in the schools for home economics, the organization of elementary schools in large cities, and the status of adult education.

The appendixes of many of the governmental publications cited above indicate sampling methods used and problems encountered. A wealth of illustrative material is available from these sources.

Use of Available Lists in Sampling Procedures

Lists of school districts from the Bureau of the Census prepared for the 1952 and 1957 Census of Governments have been the principal source for sampling of school districts. In U.S. Office of Education studies, Mason and Bain (1959) depended upon the list of school districts from the 1952 Census of Governments; Mason (1961) used the 1957 Census of Governments listing. Coon (1962) for the U.S. Office of Education, Flanagan and others (1962) for Project Talent, and the National Science Foundation (1963) utilized the basic information in the *Directory of Public Secondary Day Schools 1958-59* by Ramsey (1961). In his study of large city elementary schools, Dean (1960) relied upon the 1950 Census of Population to locate cities to be studied.

It was not until 1959, when the U.S. Office of Education began a comprehensive study of school district expenditures, that a list of school systems was specially obtained for an Office sample study. The results of this listing, which are presented in a report of the U.S. Department of Commerce, Bureau of the Census (1960b), show a reduction in school systems from 52,913 in 1957 to 42,429 in 1959. In 1952, there were 69,755 school systems in the United States; and in 1962, only 37,438. The Woodward (1961) survey of adult education statistics is a good example of information that can be obtained by sampling but is not readily derived from a regular reporting system.

The National Science Foundation (1963) employed the Register for Junior and Senior High School Science and Mathematics Teaching Personnel 1960-61, compiled by NSTA. NSTA used the basic data in the Office directory for both private and secondary schools. As the report by the Foundation indicated, cluster sampling error and additional complications in the collection of data would have been introduced without use of such a list.

The multistage sampling of the U.S. Department of Agriculture, described by Anderson and Hoofnagle (1960), for data on school food services began with the Bureau of the Census list of school districts. Lists of school plants were obtained by the Bureau of the Census from a sample of school districts. After a determination of which school plants conducted a lunch program, schools were selected on an area-probability basis. Invoices from these schools for two months were obtained. The sample was subdivided into six approximately equal subgroups of schools with invoices of deliveries taken in a rotating pattern in order to provide coverage during all months of the year. Because of sample size and of the relatively uniform nature of the National School Lunch Program, there was a smaller sampling error for schools participating in the national program than there was for nonparticipating schools.

When the first step in a sample survey entails the establishment of a universe listing of school systems or secondary schools or other units to be sampled, many of the advantages of sample surveying are lost. Establishing and maintaining a universe list is of major importance in conducting sample surveys in education. When old universe listings must be used, the nonsampling error is increased because of out-of-scope systems and other problems. Reports of most of the sample surveys previously cited have indicated such difficulties. Reorganization of school systems as evidenced by the previously indicated reduction in numbers of school districts and the continuing establishment of new schools during the past decade emphasize the need to identify and to maintain school universes.

Sampling as a Means of Identifying Information in a Census

Sampling plays an important role in the identification of information in a census. Two publications from the 1957 Census of Governments deserve mention. In a U.S. Department of Commerce, Bureau of the Census (1958) publication, detailed classes of financial information were obtained on a sampling basis. This action was motivated by the desire to obtain a maximum amount of information within a limited budget. It is far easier to control nonresponse errors for a small group of school districts than for the entire universe. The U.S. Department of Commerce, Bureau of the Census (1959) provided estimates of the market value of property in the United States by state. It would have been impossible to have obtained this information in almost any other manner. In its investigation, the U.S. Office of Education used a similar procedure in attaining almost complete universe coverage. In several studies, extremely small schools were sampled in the early 1950's.

Teacher Surveys

During the past several years, the National Education Association, Research Division (1960), along with the U.S. Office of Education, has been the primary source of information about education on the national level. Its efforts to use the power of sampling theory in sending 1,000 questionnaires to a sample of the universe of 1.3 million teachers in the United States was evidence of advance in this area. In an interesting paper, Robinson and McCall (1961) discussed response rates and variance estimates for teacher surveys conducted by the National Education Association. Their position supported what has been advocated for a decade or two on such surveys: (a) sampling to achieve coverage, (b) avoiding nonresponse, and (c) controlling uncertainty which can be related to random design.

Project Talent

Conducted by the American Institute for Research, Project Talent is an example of a large-scale study of characteristics of schools and their effects upon students—a study which exemplifies the necessity of sophisticated sample theory and technique in education. At one and the same time, it requires what the statistician calls a multiparametric model, requires a model for subpopulation studies, and represents a mixed model. Initiated in 1957, the investigation has been reported in various publications, of which the monograph by Flanagan and others (1962) is representative. In this volume, the authors pointed out a variety of parameters to which the study was directed and for which the subject of regulating the degree of uncertainty was difficult. In the investigation, 440,000 students in grades 9 through 12 in 1,353 secondary schools were tested with instruments concerned with an assessment of aptitudes, achievement, attitudes, interests, personal characteristics, biographical data, and personal plans. A stratified, random cluster sample (by school, by state, by size of senior class, and by holding power) was used primarily to obtain a "representative" sample with a degree of efficiency. Because of the large number of students in the study, it should yield much valuable information and be useful in blueprinting some of the variances with which the statistician must deal in future samples of secondary school populations on similar dimensions.

Furno's Study

An additional study which has involved sampling is significant because it is one of the few investigations which is concerned with some of the day-to-day budgeting, finance, and operational problems of school systems. Furno (1962) undertook a statistical evaluation of the Cost of Education Index, which in midyear reports such data as average amount of money spent per pupil by various expenditure accounts and by region and size of class. In using the Index, a school system can compare cost data with the national average or with the average of its class of district. Since medians and percentiles are reported by subgroups or strata, the method of sampling is one of stratification. It would be valuable if sampling variances could be reported for all subpopulations. Since school systems are notoriously inaccurate and unstandardized in reporting, a study of reporting errors (by sampling methods) would be desirable.

Limitations of Educational Applications

Many of the foregoing educational applications of survey sampling methods suffer from one or more of the following limitations:

1. Unclear objectives of a survey preventing a clear specification of the parameters needed and of the degree of precision required
2. Inefficient design from failure to take full advantage of strategies available with existing variance models
3. Inappropriate variance formula in computing confidence intervals.

A common fault is the use of a cluster sample and a simple random error variance. The study reported by Robinson and McCall (1961) showed the standard error for mean age of teachers as computed by the unrestricted random method to be .32 and by an appropriate cluster formula to be 1.85—a considerable difference.

4. A "good" probability sample, but unsuccessful coverage and lack of treatment of nonresponse and other biases

5. Sampling on the basis of an acceptable probability model, but not computing or reporting observed precision

6. Use of sophisticated sample design for which an error-loaded questionnaire or test on ambiguous or multidimensional attributes introduces psychometric errors of unknown but undoubtedly large magnitude

7. Attempts to reproduce a population or subpopulation of schools, teachers, or pupils in many dimensions accompanied by the introduction of impractical multiparametric design problems.

Bibliography

Anderson, Kenneth E., and Hoofnagle, William S. *The Market for Food in Public Schools.* U.S. Department of Agriculture, Agricultural Marketing Service, Marketing Research Report No. 377. Washington, D.C.: Government Printing Office, January 1960. 58 pp.

Cochran, William G. *Sampling Techniques.* Revised edition. New York: John Wiley & Sons, 1963. 413 pp.

Coon, Beulah I. *Home Economics in the Public Secondary Schools.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 661. Washington, D.C.: Government Printing Office, 1962. 156 pp.

Cornell, Francis G. "Sampling Methods." *Encyclopedia of Educational Research.* (Edited by Chester W. Harris.) Third edition. New York: Macmillan Co., 1960. pp. 1181-84.

Cornell, Francis G. "The Sampling Problem in Educational Research." *Research Design and Analysis: Second Annual Phi Delta Kappa Symposium on Educational Research.* (Edited by Raymond O. Collier, Jr., and Stanley M. Elam.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1961. Chapter 3, pp. 65-95.

Dalenius, Tore. "Recent Advances in Sample Survey Theory and Methods." *Annals of Mathematical Statistics* 33: 325-49; June 1962.

Davis, James A. *Great Aspirations: Volume I. Career Decisions and Educational Plans During College.* U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1194. Chicago: National Opinion Research Center, University of Chicago, 1963. 590 pp.

Dean, Stuart E. *Elementary School Administration and Organization.* U.S. Department of Health, Education, and Welfare, Office of Education, Bulletin 1960, No. 11. Washington, D.C.: Government Printing Office, 1960. 126 pp.

Deming, W. Edwards. *Sample Design in Business Research.* New York: John Wiley & Sons, 1960. 517 pp.

Flanagan, John C., and others. *Studies of the American High School.* Project Talent Monograph Series, Monograph No. 2. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 226. Pittsburgh, Pa.: University of Pittsburgh Press, 1962. 368 pp.

Furno, Orlando F. "A Statistical Look at the Cost of Education Index." *School Management* 6: 171-72, 174-78; January 1962.

Kish, Leslie. "The Design of Sample Surveys." *Research Design and Analysis: Second Annual Phi Delta Kappa Symposium on Educational Research.* (Edited by Raymond O. Collier, Jr., and Stanley M. Elam.) Bloomington, Ind.: Phi Delta Kappa (Eighth St. and Union Ave.), 1961. Chapter 2, pp. 45-63.

Mason, Ward S. *The Beginning Teacher: Status and Career Orientations.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 644. Washington, D.C.: Government Printing Office, 1961. 196 pp.

Mason, Ward S., and Bain, Robert K. *Teacher Turnover in the Public Schools 1957-58.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 608. Washington, D.C.: Government Printing Office, 1959. 38 pp.

National Education Association, Research Division. "Small-Sample Techniques." *NEA Research Bulletin* 38: 99-104; December 1960.

National Science Foundation. *Employment in Professional Mathematical Work in Industry and Government.* Washington, D.C.: Government Printing Office, 1962. 82 pp.

National Science Foundation. *Secondary School Science and Mathematics Teachers: Characteristics and Service Loads.* Washington, D.C.: Government Printing Office, 1963. 45 pp.

Ramsey, Leah W. *Directory of Public Secondary Day Schools 1958-59.* Publication No. OE-20031. Washington, D.C.: Government Printing Office, 1961. 163 pp.

Robinson, Glen, and McCall, Chester H., Jr. *Some Statistical Findings from Nationwide Teacher Polling.* Paper presented at the annual meeting of the American Statistical Association, New York City. Washington, D.C.: Research Division, National Education Association, 1961. 15 pp.

Slonim, Morris James. *Sampling in a Nutshell.* New York: Simon and Schuster, 1960. 145 pp.

U.S. Department of Commerce, Bureau of the Census. *Finances of School Districts.* 1957 Census of Governments, Vol. 3, No. 1. Washington, D.C.: Government Printing Office, 1958. 367 pp.

U.S. Department of Commerce, Bureau of the Census. *Taxable Property Values in the United States.* 1957 Census of Governments, Vol. 5. Washington, D.C.: Government Printing Office, 1959. 145 pp.

U.S. Department of Commerce, Bureau of the Census. *The Post-Enumeration Survey: 1950.* Technical Paper No. 4. Washington, D.C.: the Bureau, 1960. 99 pp. (a)

U.S. Department of Commerce, Bureau of the Census. *Public School Systems in 1960.* State and Local Government Special Studies, G-SS, No. 44. Washington, D.C.: the Bureau, 1960. 31 pp. (b)

U.S. Department of Commerce, Bureau of the Census. *School Enrollment and Educational Attainment, for the United States: 1960.* Current Population Reports—Supplementary Reports, Series PC(S1)-20. Washington, D.C.: the Bureau, 1962. 6 pp. (a)

U.S. Department of Commerce, Bureau of the Census. *School Enrollment: October 1961.* Current Population Reports—Population Characteristics, Series P-20, No. 117. Washington, D.C.: the Bureau, 1962. 11 pp. (b)

U.S. Department of Commerce, Bureau of the Census. *Summary of Governmental Finances in 1961.* Governmental Finances in the United States: 1961, Series G-GF61-No.1. Washington, D.C.: the Bureau, 1962. 25 pp. (c)

U.S. Department of Commerce, Bureau of the Census. *City Employment in 1962.* Government Employment: 1962, Series G-GE62-No. 2. Washington, D.C.: the Bureau. 1963. 15 pp. (a)

U.S. Department of Commerce, Bureau of the Census. *Estimates of Illiteracy, by States: 1960.* Current Population Reports, Series P-23, No. 8. Washington, D.C.: the Bureau, 1963. 2 pp. (b)

U.S. Department of Commerce, Bureau of the Census. *State Distribution of Public Employment in 1962.* Government Employment: 1962, Series G-GE62-No.1. Washington, D.C.: the Bureau, 1963. 23 pp. (c)

Woodward, Marthine V. *Statistics of Public School Adult Education 1958-59.* U.S. Department of Health, Education, and Welfare, Office of Education, Circular No. 660. Washington, D.C.: Government Printing Office, 1961. 49 pp.

Additional References

Anderson, T. W. "A Modification of the Sequential Probability Ratio To Reduce the Sample Size." *Annals of Mathematical Statistics* 31: 165-97; March 1960.

Bechhofer, Robert. "A Note on the Limiting Relative Efficiency of the Wald Sequential Probability Ratio Test." *Journal of the American Statistical Association* 55: 660-63; December 1960.

Coale, Ansley J., and Stephan, Frederick F. "The Case of the Indians and the Teen-Age Widows." *Journal of the American Statistical Association* 57: 338-47; June 1962.

David, Martin. "The Validity of Income Reported by a Sample of Families Who Received Welfare Assistance During 1959." *Journal of the American Statistical Association* 57: 680-85; September 1962.

Fan, Chung-Teh; Muller, Mervin E.; and Rezucha, Ivan. "Development of Sampling Plans by Using Sequential (Item by Item) Selection Techniques and Digital Computers." *Journal of the American Statistical Association* 57: 387-402; June 1962.

Fellegi, Ivan P. "Sampling with Varying Probabilities Without Replacement: Rotating and Non-Rotating Samples." *Journal of the American Statistical Association* 58: 183-201; March 1963.

Fields, R. I.; Kramer, C. Y.; and Clunies-Ross, C. W. "Joint Estimation of the Parameters of Two Normal Populations." *Journal of the American Statistical Association* 57: 446-54; June 1962.

Fiering, Myron B. "On the Use of Correlation To Augment Data." *Journal of the American Statistical Association* 57: 20-32; March 1962.

Glasser, Gerald J. "A Distribution-Free Test of Independence with a Sample of Paired Observations." *Journal of the American Statistical Association* 57: 116-33; March 1962. (a)

Glasser, Gerald J. "Variance Formulas for the Mean Difference and Coefficient of Concentration." *Journal of the American Statistical Association* 57: 648-54; September 1962. (b)

Goodman, Leo A. "Statistical Methods for the Mover-Stayer Model." *Journal of the American Statistical Association* 56: 841-68; December 1961.

Harter, H. Leon. "The Use of Sample Ranges in Setting Exact Confidence Bounds for the Standard Deviation of a Rectangular Population." *Journal of the American Statistical Association* 56: 601-609; September 1961.

Hartley, H. O., and Rao, J. N. K. "Sampling with Unequal Probabilities and Without Replacement." *Annals of Mathematical Statistics* 33: 350-74; June 1962.

Hoeffding, Wassily. "Lower Bounds for the Expected Sample Size and the Average Risk of a Sequential Procedure." *Annals of Mathematical Statistics* 31: 352-68; June 1960.

Jaeger, Carol M., and Pennock, Jean L. "An Analysis of Consistency of Response in Household Surveys." *Journal of the American Statistical Association* 56: 320-27; June 1961.

Kindahl, James K. "Estimation of Means and Totals from Finite Populations of Unknown Size." *Journal of the American Statistical Association* 57: 61-91; March 1962.

Kish, Leslie. "Efficient Allocation of a Multipurpose Sample." *Econometrica* 29: 363-85; July 1961.

Kish, Leslie; Namboodiri, N. Krishnan; and Pillai, R. Krishna. "The Ratio Bias in Surveys." *Journal of the American Statistical Association* 57: 863-76; December 1962.

Konijn, Hendrik S. "Regression Analysis in Sample Surveys." *Journal of the American Statistical Association* 57: 590-606; September 1962.

Leone, F. C.; Rutenberg, Y. H.; and Topp, C. W. "The Use of Sample Quasi-Ranges in Setting Confidence Intervals for the Population Standard Deviation." *Journal of the American Statistical Association* 56: 260-72; June 1961.

Madansky, Albert. "Tests of Homogeneity for Correlated Samples." *Journal of the American Statistical Association* 58: 97-119; March 1963.

Morgan, James N., and Sanquist, John A. "Problems in the Analysis of Survey Data, and a Proposal." *Journal of the American Statistical Association* 58: 415-34; June 1963.

Murthy, M. N., and Sethi, V. K. "Randomized Rounded-Off Multipliers in Sampling Theory." *Journal of the American Statistical Association* 56: 328-34; June 1961.

Nieto de Pascual, Jose. "Unbiased Ratio Estimators in Stratified Sampling." *Journal of the American Statistical Association* 56: 70-87; March 1961.

Raj, Des. "On Matching Lists by Samples." *Journal of the American Statistical Association* 56: 151-55; March 1961.

Robson, D. S., and Vithayasai, C. "Unbiased Componentwise Ratio Estimation." *Journal of the American Statistical Association* 56: 350-58; June 1961.

Ross, Alan. "Variance Estimates in 'Optimum' Sample Designs." *Journal of the American Statistical Association* 56: 135-42; March 1961.

Sandiford, Peter J. "A New Binomial Approximation for Use in Sampling from Finite Populations." *Journal of the American Statistical Association* 55: 718-22; December 1960.

Schaeffer, Ester B., and Dwyer, Paul S. "Computation with Multiple k-Statistics." *Journal of the American Statistical Association* 58: 120-51; March 1963.

Siegel, Sidney, and Tukey, John W. "A Nonparametric Sum of Ranks Procedure for Relative Spread in Unpaired Samples." *Journal of the American Statistical Association* 55: 429-45; September 1960.

Sukhatme, Balkrishna V. "Some Ratio-Type Estimators in Two-Phase Sampling." *Journal of the American Statistical Association* 57: 628-32; September 1962.

Weiss, Lionel. "Tests of Fit Based on the Number of Observations Falling in the Shortest Sample Spacings Determined by Earlier Observations." *Annals of Mathematical Statistics* 32: 838-45; September 1961.

Weiss, Lionel. "On Sequential Tests Which Minimize the Maximum Expected Sample Size." *Journal of the American Statistical Association* 57: 551-66; September 1962. (a)

Weiss, Lionel. "A Sequential Test of the Equality of Probabilities in a Multinomial Distribution." *Journal of the American Statistical Association* 57: 769-74; December 1962. (b)

Williams, W. H. "Generating Unbiased Ratio and Regression Estimators." *Biometrics* 17: 267-74; June 1961.

Williams, W. H. "On Two Methods of Unbiased Estimation with Auxiliary Variates." *Journal of the American Statistical Association* 57: 184-86; March 1962. (a)

Williams, W. H. "The Variance of an Estimator with Post-Stratified Weighting." *Journal of the American Statistical Association* 57: 622-27; September 1962. (b)

CHAPTER VII

Scaling and Measurement Theory

HAROLD SCHIFFMAN and SAMUEL MESSICK *

DURING THE past few years, quantitative psychology has developed from its earlier concerns with psychophysics and mental testing into what amounts to the general application of mathematics to the solution of psychological problems (Gulliksen, 1959). Several new areas have come under the purview of mathematical psychology—for example, statistical inference theories of signal detection, scaling models incorporating mathematical descriptions of individual differences, and computer simulations of cognitive processes.

Disparate fields of psychology have been coalesced through the maturation of mathematical theories, with a strong trend toward axiomatics appearing among the theoreticians in the field (Luce, Bush, and Galanter, 1963a, b). These theories extract the invariances of structure and form from domains which apparently differ on the surface. Thus, for example, one can see concepts of information theory permeating both theoretical and empirical considerations within psychophysics and scaling. One finds game theory being incorporated into signal detection models, and both of these approaches have their impact on mathematical learning theory.

These advances in theory, however, have not usually been matched by equally extensive empirical research, in either experimentation or validation. The theoretical aspects of psychometrics thus appear to have outdistanced by far the verification of empirical consequences of these theories. Considerable theoretical emphasis, for example, has been placed on the metric properties of the numbers generated by a theory—an emphasis making explicit the level of measurement (ordinal, interval, or ratio) consistent with the model. However, this survey of research indicates that neither the numbers themselves nor the transformations allowable on them are put to much further use.

Finally, it must be recognized that, unlike physics in which the measuring instruments are special cases of more general laws, measurement techniques in psychology usually have coordinate status. They frequently imply psychological theories of the same order as the hypotheses they are used to evaluate. Therefore, when these measurement techniques are interposed between a hypothesis and the raw data, their filtering action may only allow information conformable with the axioms of the measuring device to become visible to the experimenter (Messick and Ross, 1962).

---

* Parts of this chapter were completed while the second cited author was a fellow at the Center for Advanced Study in the Behavioral Sciences. The authors wish to thank Herbert Gerjuoy and John Keats for their comments on the manuscript.

History of Psychometrics

Psychometrics generally marks its beginning with the publication of Fechner's *Elemente der Psychophysik*, and *Psychometrika* devoted a commemorative issue jointly to the one-hundredth anniversary of this event and to the twenty-fifth anniversary of the founding of the Psychometric Society. In this commemorative volume, a series of papers on various aspects of the history of psychometrics appeared. Dunlap (1961) critically evaluated the status of quantitative psychology and argued with great force that present-day psychometrics is "extremely narrow in its concept of what constitutes the development of quantitative rationale for the solution of psychological problems" (p. 69) and that it has "become complacent, perhaps authoritarian, and unimaginative in . . . approach . . . " (p. 70). Boring (1961) commented upon some contributions of Fechner as the "inadvertent founder of psychophysics." Guilford (1961) discussed trends in psychometric publications and reviewed some active issues in the field. Gulliksen (1961) provided a substantive history of scaling and discussed the laws of comparative and categorical judgment, multidimensional scaling, and other recent developments in measurement in the context of establishing scientific laws of human behavior. He hoped, in conclusion, that these methods would not be considered ends in themselves, but would be used to investigate the properties of judgments and their functional relations with other variables.

An informative historical introduction to psychometrics also appeared in a scholarly article by Ramul (1960). From his concrete illustrations of eighteenth century scientists trying to come to grips with metrical methods for dealing with human beings, one can appreciate the difficulties encountered when the concepts and categories of thought are not sufficiently definite for such an approach.

Psychophysics

It is well documented that measures of sensory magnitude obtained by category rating and direct magnitude estimation methods (Stevens, 1958, 1960) produce two different scales that are nonlinearly related to each other (Stevens and Galanter, 1957). Galanter and Messick (1961), investigating this problem experimentally, found that a processed category (successive intervals) scale was a logarithmic function of the magnitude scale. Rationales for the logarithmic relation have been advanced by Eisler (1963), by Helm, Messick, and Tucker (1961), and by others.

The anniversary issue of *Psychometrika* also included a discussion of recent developments in psychophysics. Keith Smith (1961) discussed certain issues in the presentation of stimulus material to the subject, which he calls "stimulus programming." Swets (1961a) reviewed the theory of signal detectability and discussed its implications for psychophysics. Stevens (1961) continued his long-standing controversy with the Fechner-Thurstone approach and concluded that ratio scales arrived at by direct methods "have greater demonstrable usefulness . . . than any Fechnerian scale yet devised" (p. 46).

However, in a thoughtful article on some of Stevens' assumptions in psychophysics, Warren and Warren (1963) considered the problem of whether quantitative responses about stimuli do or do not directly reflect the subject's estimates of stimulus intensity. They offered, as an alternative to Stevens' view, a physical correlate theory in which psychophysical judgments are considered to be estimates of physical magnitudes associated with varying degrees of stimulation. Loudness judgments, for example, would be based upon experience with distance judgments in spatial localization (Warren, 1963). In support of this theory, they found that judgments of half loudness and half brightness were each equivalent to judgments of doubling the distance from the source. If true, this theory has powerful implications that would necessitate a re-examination of psychophysical procedures and analysis.

Signal Detection

In psychophysical experiments, the subjects' responses to identical stimuli vary, and this variation occurs systematically as the experiment proceeds. A consideration of this learning aspect in repeated psychophysical judgments plus the observation that payoff functions are important in determining thresholds are the major points of departure for theories of signal detection (Goldiamond, 1962; Green, 1960; Swets, 1961a).

Signal detection theory was developed from statistical decision theory in an attempt to specify optimal behavior in situations requiring a choice between alternative hypotheses on the basis of an observed event (Swets, Tanner, and Birdsall, 1961). Technical procedures were then developed to describe the behavior of subjects attempting to discriminate patterns with both information and noise from those patterns which contain only noise. This decision theory approach generalizes the conventional notion of a threshold as a fixed limit to the concept of a criterion range of acceptance (Atkinson, 1963; Swets, 1961b). A further advance in this formulation over the classical psychophysical methods is the insistence that during the course of an experiment the subject learns and that this learning will affect the probabilities of choice (Luce, 1963).

Information Theory

Since its continuing development by Fisher, Shannon, and others, information theory has had an impact on psychology which goes far beyond the relatively small number of studies attempting to incorporate and to test certain of its consequences.

Fortunately, a book is available which provides an extensive frame of reference for considering information theory in relation to psychological theory and experimentation. This book, *Uncertainty and Structure as Psychological Concepts*, by Garner (1962), is the outcome of several years work on the uses of information theory in psychology. It provides an introduction to the mathematical basis of information and considers problems and experiments in univariate, bivariate, and multidimensional uncertainties. Empirical research is reviewed in each section and is placed in a theoretical context; examples are redundancy in language, concept formation, coding, and category scaling.

Since the meaning of simple stimuli is related to the uncertainty of the stimulus pattern, it can be handled by informational analysis. In this connection, Alluisi (1960) reviewed recent studies of form perception and explicated the usages of information measures in perceptual studies.

A further extension of the use of information concepts into the statistical methods of psychology was given by Ross (1962), who derived an analytic relationship in the unidimensional case between an information measure and variance. He then developed a recipe for finding the relation between information and the *n*-dimensional analogue of variance. Having developed this generalized relation, he extended the argument by noting that the latent roots of a correlation matrix are related to variance and hence to information, and he thereby implicated factor analysis with information theory. The number and relative sizes of the roots were found to be related in a simple manner to the amount of pure information in a factor structure.

Multidimensional Scaling

The psychological processes involved in making the responses required by a mathematical theory are usually not emphasized in discussing the models. In the December 1957 REVIEW, Messick and Abelson provided a tentative classification of certain scaling theories in terms of their implicit psychological models of what the subject is "really" doing. They pointed out (a) that deterministic models assumed fixed "real" scales of preference or judgment, (b) that stochastic models implied that the subject acted as if he consulted a table of random numbers before making choices or that he shifted his standards of judgment through time or that since he was confused his momentary impression was subject to random fluctuation, and (c) that certain noncommital models did not take a stand with respect to an underlying psychological explanation.

On the other hand, Shepard (1962a, b) introduced a generalized model for multidimensional scaling with an explicit consideration of the psychological processes involved in several disparate tasks. From this consideration of the properties of a wide variety of tasks (Shepard, 1960), Shepard (1962a) concluded that in each case the subject's judgment or response was a function of proximity among the stimuli. Shepard then considered the intuitive isomorphisms that are expected to hold between the psychological idea of proximity and its mathematical representation. For example, the notion that A is close to B implies that B is close to A, given the metric requirement that the distance between two stimuli be symmetric. Since it is by no means certain that proximity measures are related to distance measures by a single fixed rule for all domains, Shepard considered the general case of multidimensional scaling where distances are required only to be monotonic functions of proximities.

Thus, in Shepard's (1962a, b) formulation of multidimensional scaling there is no specified distance function; only a monotonic relation between proximity and distance is kept invariant. The objectives of the analysis are to determine (a) the minimum number of dimensions of a Euclidean space such that distances in the space are monotonically related to the initial proximity measures, (b) a set of orthogonal coordinates for the points in the minimum space, and (c) a plot showing the form of the relation between the initially unknown distances and the experimental proximity measures. It is important to note that only two conditions, minimum dimensionality and invariant monotonicity, lead to a unique solution. A computer program was also developed (Shepard, 1962a, b) to implement the numerical methods of the analysis. Finally, Shepard also considered the implications of this technique for factor analysis, standard multidimensional scaling, and ordered metric scaling.

A further elaboration of Shepard's work was undertaken by Kruskal (1963), who pointed out that Shepard's technique still lacked a solid statistical foundation—that is, an explicit mathematical-statistical definition of how well the data have been compressed under the criterion of monotone invariance. Kruskal developed a theoretical rationale and practical procedure for attacking this problem. He defined a quantitative measure of fit called stress (a measure of how much the theoretical configuration differed from the raw data) and developed a computer program to find the configuration that would essentially minimize the stress. This program also permits the use of certain non-Euclidean distance functions, such as Minkowski *r*-metrics.

Guttman (undated) independently considered the problem of enlarging the class of distance measures. It is of interest that, although starting from a set of presuppositions completely different from Shepard's, he came out with essentially the procedures stated above. Guttman was inhibited from coming to a full feasible solution by the lack of high-speed computer facilities.

Another approach to the derivation of a multidimensional configuration from ordinal distance information was attempted in terms of a generalization of Coombs' unfolding technique from its unidimensional basis to the multidimensional case (Bennett and Hays, 1960; Hays and Bennett, 1961; McElwain and Keats, 1961).

Ekman (1963) generalized direct ratio scaling methods into a multidimensional model and suggested a procedure in which the observer judges the proportion of a standard percept that each stimulus percept contains. The model represents these judgments as projections on unit reference vectors, from which scalar products are derived as a basis for factor analysis.

Applications

Multidimensional scaling has been applied in widely different areas during the past three years, including studies of emotions in faces (Abelson and Sermat, 1962), of the perception of political leaders (Messick, 1961), and of the social dimensions of a jury table (Strodtbeck and Hook, 1961). Wilson (1963) performed a multidimensional analysis of confusion errors (Shepard, 1957) in an attempt to identify cues used in the recognition of English consonants, and Keats (1962) used multidimensional unfolding to compare preferences for artists, heroes, saints, and scholars in Malaya and Australia.

One difficulty in applying standard multidimensional scaling procedures is the enormous number of judgments required from each subject. One approach to reducing this large number of judgments was suggested by Gulliksen and Tucker (1961) for unidimensional scaling. Their procedure utilized multiple rank order judgments obtained in balanced incomplete block designs as a basis for deriving paired comparison information.

Individual Differences

Historically, two major approaches have been followed in analyzing scaling data obtained from groups of individuals. One approach has been to compute group averages and to generalize the findings "on the average." The other procedure has been to analyze each person's responses separately and to tabulate the results individual by individual. Although the first method is straightforward and frequently used, it is possible that the averaging of responses across subjects may mask consistent individual differences in reaction to the stimuli and thus may produce a conglomerate result that may not adequately summarize the individual data (Messick, 1961). The second method of treating each individual separately also presents difficulties, particularly in describing the results for groups of individuals and in comparing the results for several individuals and groups. A third alternative has recently been suggested—an approach that both isolates consistent individual differences in response to stimuli and derives separate scales for each consistent viewpoint (e.g., Hays and

Bennett, 1961). Models and procedures for determining dimensions of individual differences have been presented for unidimensional scaling by Tucker (1960) and Slater (1960), and an individual differences model for multidimensional scaling was described by Tucker and Messick (1963). Helm and Tucker (1962) applied this latter model to the multidimensional scaling of colors and isolated separate color spaces for corresponding individual differences in color deficiency.

Computer Models

A relatively new approach to model building using computers and computer languages has begun to bear fruit in recent years. We have seen in an earlier section how computer models for multidimensional scaling with an unknown distance function (Kruskal, 1963; Shepard, 1962a, b) have made feasible the tedious numerical analyses required for a solution. It is hoped that similar computer formulations will also prove useful in other areas for the derivation of complex measurement parameters.

Because the computer permits a massive increase in the complexity of the relationships that can be manipulated and also permits dynamic integration over time, computer models offer a powerful adjunct to mathematical models in areas involving complex cognitive and personality processes (Borko, 1962; Green, 1961, 1963; Newell and Simon, 1963; Tomkins and Messick, 1963).

Reviews and Integrative Summaries

Developments from 1960 to 1962 in scaling, test theory, and factor analysis were reviewed by Tucker (1963). Recent contributions in scaling, including applications in psychophysics, test theory, and the measurement of utility, were presented in *Psychological Scaling: Theory and Applications* (Gulliksen and Messick, 1960). Restle (1961) developed a mathematical theory of choice and discussed its application in areas of utility measurement, recognition, detection, psychophysics, and response times.

Luce, Bush, and Galanter (1963a) edited an impressive three-volume *Handbook of Mathematical Psychology*, of which the first volume has appeared. This handbook, written by and for professionals in the field, provides both a summary and an analysis of the major mathematical theories in psychology. By a careful, and frequently formal and axiomatic, presentation of the theoretical structures under consideration, this volume serves to explicate the assumptions underlying mathematical psychology in general as well as psychological measurement in particular. The presentation, which is maintained at a uniformly high level, offers an immense compression of material while keeping the essentials articulated. The topics covered in the first volume include basic measurement theory, choice experiments, detection and recognition, discrimination, psychophysical scaling, stochastic latency mechanisms, estimation and evaluation, and computers. Luce, Bush, and Galanter (1963b) have also edited a two-volume companion book of convenient source material for the researcher and teacher of advanced psychometrics.

Mathematics and Psychology

In conclusion, the writers would like to emphasize a point bearing upon the presumptions of mathematical psychology. Many of the existing mathematical theories of psychology have been constructed as formal systems and have made explicit their mathematical properties and the conditions required for their application. However, these formulations are essentially theories of human processes cast in mathematical terms, and the role of theory in directing and even controlling research is well known. Limitations in current psychological views of human processes, built into the mathematical theory as axiom and supposition, will still be present in the transformed notation, along with additional constraints that are associated with the structure of the superimposed mathematical system. These limitations will directly influence the form that subsequent empirical research takes, and this influence, which is pervasive, extends from the manner in which data are gathered to the choice of topics deemed important.

For formal theories—mathematical, statistical, or logical—to become empirical theories they must undergo a process whereby the elements and rules of the formal model are interpreted and given empirical meaning. Propositions true in a formal system by implication also need to be true in an empirical sense after interpretation. This most important topic of investigating the psychological meaning and relevance of the axioms has not yet received in any measure the systematic effort so far invested in formulating the axioms.

Bibliography

Abelson, Robert P., and Sermat, Vello. "Multidimensional Scaling of Facial Expressions." *Journal of Experimental Psychology* 63: 546-54; June 1962.

Alluisi, Earl A. "On the Use of Information Measures in Studies of Form Perception." *Perceptual and Motor Skills* 11: 195-203; October 1960.

Atkinson, Richard C. "A Variable Sensitivity Theory of Signal Detection." *Psychological Review* 70: 91-106; January 1963.

Bennett, Joseph F., and Hays, William L. "Multidimensional Unfolding: Determining the Dimensionality of Ranked Preference Data." *Psychometrika* 25: 27-43; March 1960.

Boring, Edwin G. "Fechner: Inadvertent Founder of Psychophysics." *Psychometrika* 26: 3-8; March 1961.

Borko, Harold, editor. *Computer Applications in the Behavioral Sciences.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 633 pp.

Dunlap, Jack W. "Psychometrics—A Special Case of the Brahman Theory." *Psychometrika* 26: 65-71; March 1961.

Eisler, Hannes. "Magnitude Scales, Category Scales, and Fechnerian Integration." *Psychological Review* 70: 243-53; May 1963.

Ekman, Gösta. "A Direct Method for Multidimensional Ratio Scaling." *Psychometrika* 28: 33-41; March 1963.

Galanter, Eugene, and Messick, Samuel. "The Relation Between Category and Magnitude Scales of Loudness." *Psychological Review* 68: 363-72; November 1961.

Garner, Wendell R. *Uncertainty and Structure as Psychological Concepts.* New York: John Wiley & Sons, 1962. 369 pp.

Goldiamond, Israel. "Perception." *Experimental Foundations of Clinical Psychology.* (Edited by Arthur J. Bachrach.) New York: Basic Books, 1962. Chapter 9, pp. 280-340.

Green, Bert F., Jr. "Computer Models of Cognitive Processes." *Psychometrika* 26: 85-91; March 1961.

Green, Bert F., Jr. *Digital Computers in Research.* New York: McGraw-Hill Book Co., 1963. 333 pp.

Green, David M. "Psychoacoustics and Detection Theory." *Journal of the Acoustical Society of America* 32: 1189-1203; October 1960.

Guilford, J. P. "Psychological Measurement a Hundred and Twenty-Five Years Later." *Psychometrika* 26: 109-27; March 1961.

Gulliksen, Harold. "Mathematical Solutions for Psychological Problems." *American Scientist* 47: 178-201; June 1959.

Gulliksen, Harold. "Linear and Multidimensional Scaling." *Psychometrika* 26: 9-25; March 1961.

Gulliksen, Harold, and Messick, Samuel, editors. *Psychological Scaling: Theory and Applications.* New York: John Wiley & Sons, 1960. 211 pp.

Gulliksen, Harold, and Tucker, Ledyard R. "A General Procedure for Obtaining Paired Comparisons from Multiple Rank Orders." *Psychometrika* 26: 173-83; June 1961.

Guttman, Louis. *The Smallest Euclidean Space for a Configuration of Points.* Jerusalem: Israel Institute of Applied Social Research. 16 pp.

Hays, William L., and Bennett, Joseph F. "Multidimensional Unfolding: Determining Configuration from Complete Rank Order Preference Data." *Psychometrika* 26: 221-38; June 1961.

Helm, Carl E., and Tucker, Ledyard R. "Individual Differences in the Structure of Color Perception." *American Journal of Psychology* 75: 437-44; September 1962.

Helm, Carl E.; Messick, Samuel; and Tucker, Ledyard R. "Psychological Models for Relating Discrimination and Magnitude Estimation Scales." *Psychological Review* 68: 167-77; May 1961.

Keats, J. A. "Attitudes Towards Idealized Types in Australia and Malaya." *Journal of Social Psychology* 57: 353-62; August 1962.

Keith Smith, J. E. "Stimulus Programming in Psychophysics." *Psychometrika* 26: 27-33; March 1961.

Kruskal, J. B. *A Systematic Approach to Multidimensional Scaling.* Murray Hill, N.J.: Bell Telephone Laboratories, 1963. 53 pp.

Luce, R. Duncan. "A Threshold Theory for Simple Detection Experiments." *Psychological Review* 70: 61-79; January 1963.

Luce, R. Duncan; Bush, Robert R.; and Galanter, Eugene H., editors. *Handbook of Mathematical Psychology, Vol. 1.* New York: John Wiley & Sons, 1963. 491 pp. (a)

Luce, R. Duncan; Bush, Robert R.; and Galanter, Eugene, editors. *Readings in Mathematical Psychology, Vol. 1.* New York: John Wiley & Sons, 1963. 535 pp. (b)

McElwain, D. W., and Keats, J. A. "Multidimensional Unfolding: Some Geometrical Solutions." *Psychometrika* 26: 325-32; September 1961.

Messick, Samuel. "The Perceived Structure of Political Relationships." *Sociometry* 24: 270-78; September 1961.

Messick, Samuel J., and Ross, John, editors. *Measurement in Personality and Cognition.* New York: John Wiley & Sons, 1962. 334 pp.

Newell, Allen, and Simon, Herbert A. "Computers in Psychology." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Chapter 7, pp. 361-428.

Ramul, Konstantin. "The Problem of Measurement in the Psychology of the Eighteenth Century." *American Psychologist* 15: 256-65; April 1960.

Restle, Frank. *Psychology of Judgment and Choice.* New York: John Wiley & Sons, 1961. 235 pp.

ROSS, JOHN. "Informational Coverage and Correlational Analysis." *Psychometrika* 27: 297-306; September 1962.

SHEPARD, ROGER N. "Stimulus and Response Generalization: A Stochastic Model Relating Generalization to Distance in Psychological Space." *Psychometrika* 22: 325-45; December 1957.

SHEPARD, ROGER N. "Similarity of Stimuli and Metric Properties of Behavioral Data." *Psychological Scaling: Theory and Applications.* (Edited by Harold Gulliksen and Samuel Messick.) New York: John Wiley & Sons, 1960. Chapter 4, pp. 33-43.

SHEPARD, ROGER N. "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. I." *Psychometrika* 27: 125-40; June 1962. (a)

SHEPARD, ROGER N. "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. II." *Psychometrika* 27: 219-46; September 1962. (b)

SLATER, PATRICK. "The Analysis of Personal Preferences." *British Journal of Statistical Psychology* 13: 119-35; November 1960.

STEVENS, S. S. "Problems and Methods of Psychophysics.' *Psychological Bulletin* 55: 177-96; July 1958.

STEVENS, S. SMITH. "Ratio Scales, Partition Scales and Confusion Scales." *Psychological Scaling: Theory and Applications.* (Edited by Harold Gulliksen and Samuel Messick.) New York: John Wiley & Sons, 1960. Chapter 6, pp. 49-66.

STEVENS, S. S. "Toward a Resolution of the Fechner-Thurstone Legacy." *Psychometrika* 26: 35-47; March 1961.

STEVENS, S. S., and GALANTER, E. H. "Ratio Scales and Category Scales for a Dozen Perceptual Continua." *Journal of Experimental Psychology* 54: 377-411; December 1957.

STRODTBECK, FRED L., and HOOK, L. HARMON. "The Social Dimensions of a Twelve-Man Jury Table." *Sociometry* 24: 397-415; December 1961.

SWETS, JOHN A. "Detection Theory and Psychophysics: A Review." *Psychometrika* 26: 49-63; March 1961. (a)

SWETS, JOHN A. "Is There a Sensory Threshold?" *Science* 134: 168-77; July 21, 1961. (b)

SWETS, JOHN A.; TANNER, WILSON P., JR.; and BIRDSALL, THEODORE G. "Decision Processes in Perception." *Psychological Review* 68: 301-40; September 1961.

TOMKINS, SILVAN S., and MESSICK, SAMUEL, editors. *Computer Simulation of Personality: Frontier of Psychological Theory.* New York: John Wiley & Sons, 1963. 325 pp.

TUCKER, LEDYARD R. "Intra-Individual and Inter-Individual Multidimensionality." *Psychological Scaling: Theory and Applications.* (Edited by Harold Gulliksen and Samuel Messick.) New York: John Wiley & Sons, 1960. Chapter 13, pp. 155-67.

TUCKER, LEDYARD R. "Scaling and Test Theory." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1963. Vol. 14, pp. 351-64.

TUCKER, LEDYARD R., and MESSICK, SAMUEL. *An Individual Differences Model for Multidimensional Scaling.* Research Bulletin 63-13. Princeton, N.J.: Educational Testing Service, 1963. 53 pp.

WARREN, RICHARD M. "Are Loudness Judgments Based on Distance Estimates?" *Journal of the Acoustical Society of America* 35: 613-14; April 1963.

WARREN, RICHARD M., and WARREN, ROSLYN P. "A Critique of S. S. Stevens' 'New Psychophysics.'" *Perceptual and Motor Skills* 16: 797-810; June 1963.

WILSON, KELLOGG V. "Multidimensional Analyses of Confusions of English Consonants." *American Journal of Psychology* 76: 89-95; March 1963.

CHAPTER VIII

Statistical Learning Models

RONALD A. WEITZMAN

THE STATISTICAL models represent perhaps the newest and fastest-growing field in psychology. It would be fatuous if not impossible to attempt to trace the growth of this field by means of a graph showing the number of articles and books published annually since 1950, when W. K. Estes turned the first patch of what has since proved to be very fertile soil. Considering together the facts that in 1950 there was only one published article and that in the past three years, the period under review, there have been so many articles that it has been difficult to select the most significant 40 (the editorial limit) for discussion is sufficient to give an idea of the pace at which the field has been developing.

Originally, the application of statistical models in psychology was exclusively in the field of learning. Although most of the recent work has continued to be done in this field, there has been a considerable body of work done in the past few years which represents an extension of the basic ideas of the learning models to other fields of psychology. When the models were confined to the study of learning, they did not have the status of a field. Now, however, because of the extensions to other fields, the models themselves must be considered a field. Indeed, a text devoted exclusively to statistical models in psychology must appear before long.

The statistical models are a variety of the mathematical models which, since the days of Fechner and Ebbinghaus, have been at least somewhat familiar to most psychologists. What distinguishes the statistical models is not only that the dependent variable in their observation equations is a statistic (either a mean, a variance, a correlation, or a proportion) but also that the assumptions on which they are based uniformly include at least one statement involving probability. However, the independent variables do not fall so neatly into one category as do the dependent variables, for, in practice, they have included not only trial numbers, random variables representing correct and incorrect responses, and proportions of reinforcements (in partial-reinforcement experiments) but also unfitted parameters.

Recently, the development of procedures for evaluating theoretically a variety of statistics has been of major concern to modelists. Earlier, interest had been confined mainly to the theoretical proportion of correct responses made by a group of subjects on each trial and especially on each of the later (asymptotic) trials of a learning experiment. Since these statistics, however, have failed to discriminate among the numerous learning models which have been presented, other statistics have been sought and considered. Among these are mean numbers of runs of correct (or incorrect) responses of various lengths, the proportion of correct (or incorrect) responses following a reinforced (or nonreinforced) trial, the correlation between random variables representing responses separated by a certain number of trials, and the variance of the total number of correct (or incorrect) responses made by all the subjects in an experiment. Some of these statistics can be applied to individual subjects as well as to groups of subjects, and some have even been used to test for individual differences.

When a number of theoretically evaluated statistics are available for several models, comparisons of these statistics with those computed directly from data afford a means for discriminating among the models and also for assessing the strengths and weaknesses of each.

Another recent development has been of a psychological as well as a statistical nature. Most psychologists know that, whereas the Bush and Mosteller models which are derived basically from mathematical notions are psychological only in application, the Estes models represent mathematical formulations and derivations of basically psychological ideas. These psychological ideas have been mainly of a Guthrian-Hullian nature. In the past three or four years, however, the Estes group of modelists has moved more strongly than previously from the Hullian to the Guthrian position. Indeed, at least one Estes modelist (Restle) has become frankly cognitive, and both he and Estes have in effect renounced previous theoretical positions. In each case, both theoretical arguments (psychological and mathematical) and empirical results appear to have been responsible for the change.

In addition to the statistical models furnished by Estes and by Bush and Mosteller, certain nonstatistical learning models have appeared from time to time. Some of these are empirically based, and some, theoretically based; among the theoretically based, some are cognitive, and some are response-strength models. One of these models will be discussed briefly in the latter part of this chapter. The major portion of this chapter, however, will deal with alterations, empirical tests, and extensions of the Estes and Bush-Mosteller models.

Alterations

The Pattern Model

In a series of papers dating from a paper by Rock in 1957, evidence has been brought to bear once again on the old continuity-versus-noncontinuity issue which Krechevsky and Spence in the 1930's and early 1940's brought so forcefully to the attention of psychologists. Statistical learning theorists have become interested in the resurrection of this apparently resolved (in favor of the continuity position) issue, and an important part of their motivation has been theoretical. The original Estes models, though based on the assumption that learning occurs in an all-or-none (noncontinuous) manner, had been able to account for the appearance of continuous learning by a device similar to that which Guthrie had used earlier in similar circumstances: an S-R connection is formed or is not formed on an experimental trial, but many such connections may be necessary to produce the response that the experimenter considers correct. In fact, Estes had identified the probability of a correct response on an experimental trial with the proportion of stimulus elements conditioned to that response on that trial. Many varieties of model based on this fundamental identification had been devised, but all suffered from the same theoretical weakness: they yielded explicit formulas for very few statistics (an embarrassing situation for a statistical model).

In 1959, Estes introduced the pattern model to psychological literature and indicated that this model had the power of producing explicit formulas for many statistics. Bower (1961) proved this when he presented explicit formulas for 22 different statistics derived from the pattern model! The pattern model, however, precludes the former identification of proportion of conditioned stimulus elements with probability of a response because it is based on the assumption that only one stimulus element (pattern) is present on each experimental trial. The gauntlet has been thrown. Either the pattern model is wrong, or learning is of a demonstrably all-or-none nature.

Recent evidence in favor of all-or-none learning has come in a variety of forms. One of the most interesting is the so-called miniature experiment (Estes, 1960)—an experiment consisting of just three, four, or five trials, of which only the first involves reinforcement (the remaining trials being "test" trials). Results of such experiments favoring the noncontinuity view were reported by Estes (1960, 1961) and by Estes, Hopkins, and Crothers (1960). Jones (1962) described results of the same kind of experiment which have, however, been only partly favorable to this view.

A significant conflict seems to be developing over another line of evidence. Bower (1961, 1962) and Trabasso (1963) showed that the average number (or proportion) of errors remains constant over trials before the last error (over presolution trials). These results are clearly in favor of the noncontinuity view. However, after presenting similar evidence, Suppes and Ginsberg (1963) indicated that when learning curves were corrected for individual differences by Vincentizing, the proportion of presolution-period errors no longer remained constant. Trabasso (1963), meanwhile, found that after making a similar correction by the method of backwards learning curves, this proportion still remained constant. Perhaps Trabasso's subjects were more homogeneous than were those of Suppes and Ginsberg, who included not only adults but also children and rats in their study and who explicitly presented evidence for individual differences among the children and the rats. Even so, Trabasso's result stands against that of Suppes and Ginsberg, and the cause of the discrepancy remains to be found.

Additional References: Peterson and others (1962); Reynolds (1963); Simon (1962); Waugh and Smith (1962); Williams (1962).

A Generalization of the Pattern Model

The pattern model, which involves only one parameter, was generalized by Suppes and Atkinson (1960) to allow for a number of parameters. In the original simple pattern model, the single parameter represents the probability that a reinforcement on any trial will be effective. In the generalized version of this model, a different parameter is possible for each response-reinforcement situation. For example, Suppes and Schlag-Rey (1962) presented, among other generalized pattern models, a model that involves, for trials following two reinforcements, a parameter that differs from the themselves-differing parameters for trials following only one or zero reinforcements, respectively, on the preceding two trials.

An Incremental Model

Atkinson (1961) marched off briefly in the opposite direction and suggested an incremental model wherein a stimulus element may be in one of a finite number of conditioning states on any experimental trial. The maximum number of states possible for an element in this model is assumed to be a function of motivation or reinforcement, and the probability of a correct response on any experimental trial is identified with the ratio of the state number of the stimulus element present on that trail to the highest state number possible for that element. Application of a single-stimulus version of this model to data involving different numbers of possible responses has shown that the maximum number of states varies with the number of possible responses and so is not, as originally hoped, simply a function of motivation or reinforcement.

A Strategy Model

Restle (1961, 1962) presented a strategy model which represents a mathematical rendition of the cognitive version of the noncontinuity view which Krechevsky presented a quarter of a century earlier. This model resembles the simple pattern model in that one element is assumed to be sampled on any experimental trial and in that if this element is already conditioned a correct response will occur. [It should be noted that Restle (1962) showed that the model is unchanged if the assumption of one element is relaxed.] Important differences between this and the pattern model, however, are (a) the interpretation of the elements available for sampling, which here are strategies (hypotheses) rather than stimulus patterns, and (b) the number of parameters, which here are three and which represent the proportions, respectively, of correct, wrong, and irrelevant strategies possible in a particular experimental situation. On any experimental trial, a subject is assumed to sample a strategy randomly from the population of possible strategies; if he samples one of the correct strategies, he will respond correctly; if he samples one of the wrong strategies, he will respond incorrectly; and if he samples one of the irrelevant strategies, he may or may not (by chance) respond correctly. Moreover, once a correct strategy has been sampled, and the response made in accordance with it is reinforced, the subject will continue to act only on the basis of that strategy. Restle (1962) maintained that this model is superior to his previous adaptation model on logical grounds; he demonstrated that in the concept identification experiments where the older model had been applied, it can account for the older model's successes and perhaps succeed where the older model failed.

The successes of the adaptation model have usually involved the principle of *additivity of cues*. Trabasso (1960), for example, showed that in using the sum of parameters estimated separately in two experiments, each involving a single relevant cue, he could successfully predict (by way of Restle's adaptation model) the number of errors made in an experiment in which both these (and only these) cues were relevant. An example of the failures was reported by Levine (1962), who showed that prolonged random reinforcement during pretraining, which according to the adaptation model should result in complete adaptation of all cues presented, does not have this effect; in fact, subjects are able to identify concepts correctly on the basis of cues which should have been rendered functionless (which should have been adapted out) during previous random reinforcement of them.

Additional References: Bourne and Bunderson (1963); Estes and Hopkins (1961); Pishkin (1960); Richards (1962).

Empirical Tests

Generally speaking, empirical tests of the models have been rather hard on them. This situation exists despite the fact that the immediate impression received on reading any single major article is usually the contrary one: namely, that the model fits the data quite well. The reason for this discrepancy seems to be that a model's originator reports (honestly) favorable results, whereas others only later adduce negative evidence. The time lag is such, however, that at any given moment it has seemed that at least one modelist has indeed cracked the code of the data.

The Matching Rule

Feldman and Newell (1961) showed that a variety of models can predict the matching of the asymptotic proportion of a response with the proportion of times that it has been reinforced in the course of an experiment and thus concluded that the occurrence or nonoccurrence of matching is not evidence exclusively in favor of any one model. The evidence regarding the appearance of matching, moreover, has always been mixed.

Additional References: Bourne (1963); Chapman (1961); Derks (1962); Greeno (1962); Hickson (1961); McCracken, Osterhout, and Voss (1962); Nies (1962); Shaffer (1962, 1963); Witte (1961).

Conditional Dependencies

Of continued interest is the proportion of correct (or incorrect) responses on a trial following the occurrence of reinforcement (or nonreinforcement) on the preceding trial or trials (event dependencies) or following the occurrence of a correct (or incorrect) response on the preceding trial or trials (response dependencies). Classifying dependencies not only as event and response but also as absolute (for example, in the case of an absolute event, dependency there being zero, one, or two reinforcements on the preceding two trials) and relative (for example, in the case of a relative event, dependency there being two possible perseverations—reinforcement-reinforcement and nonreinforcement-nonreinforcement—and two possible alternations of reinforcements and nonreinforcements on the preceding two trials—reinforcement-nonreinforcement and nonreinforcement-reinforcement), Suppes and Schlag-Rey (1962) used generalized pattern models to analyze the asymptotic data of a two-choice learning experiment. They concluded that response dependencies are more effective than event dependencies and that relative dependencies are more effective than absolute dependencies.

Additional References: Anderson and Whalen (1960); Chapman (1961); Derks (1963); Edwards (1961); Prokasy (1961); Witte (1961).

A Variety of Statistics

On a number of occasions, a variety of statistics has been presented, sometimes to evaluate a single model against data and sometimes to compare the relative fits to a single set of data of two or more models. Since not all models readily yield a large number of statistics in the form of explicit formulas, resort has occasionally been made to Monte Carlo procedures for computing statistics. Flood (1962) recommended and illustrated the use of Monte Carlo procedures on a high-speed computer as a general method for testing and comparing models. Bower (1961), however, was able (as noted earlier) to derive explicit formulas for 22 statistics from the assumptions of the pattern model and, in illustration of their use, showed that 17 out of 19 of them fitted the data of a paired-associates experiment with college student subjects better than the corresponding statistics of a Bush-Mosteller model (the only other model for which such a large number of statistics had been derived in explicit form). Modelists consider that the power of a model to yield explicit formulas for statistics is an important property, and on the basis of such a criterion the pattern model must be rated very high indeed.

What appears to be the greatest success of any statistical learning model in accounting for the data of an experiment was reported by Bower (1962) for the pattern model. From the data of a paired-associates learning experiment involving a number of groups of college students, 24 statistics were computed, each by theoretical and empirical formulas. Although the same parameter value was used in each of the theoretical formulas, all 24 comparisons were numerically close.

These results of Bower's are so startling that they immediately occasion this question: Is there not some artifact or special condition which accounts for such good results? There does appear to be one. In the same paper, Bower reported that the variance of total errors for at least some of the groups was so low that it precluded the possibility of individual differences in parameter value. Since the single parameter of the pattern model represents learning rate, this result gives rise to the suspicion that the tasks for the various groups were uniformly so easy that such differences in intelligence as might exist among the homogeneous population of college students did not extend downward sufficiently to be reflected in differences in the parameter value.

This suspicion has been confirmed by Suppes and Ginsberg (1963), who, in applying the same model to the data of concept-formation experiments with children, found not only that the children's and college students' parameter values differed but also that there were individual differences among the parameter values of the children themselves. Had Bower carried out his experiment with children or under conditions of partial rather than continuous reinforcement, almost certainly his results would not have been so good. The generality of Bower's results is thus questionable, and at present there seems to be no way of testing the pattern or any other model so extensively under conditions wherein individual differences are expected to obtain.

Extensions

Basic ideas of statistical learning models have been used in models which describe psychological phenomena not only outside the field of learning but also in areas within the field of learning—areas in which the older models do not directly apply.

Social Psychology

Suppes and Atkinson (1960) made an important extension of the pattern model to describe the behavior that occurs in a number of versions of games for two or more persons. These games are formalized social interactions, many varieties of which have been studied extensively by economic-game theorists, who suppose that in such situations human beings will behave in a strictly rational manner. Psychologists have long thought this supposition generally incorrect; and in the experiments reported by Suppes and Atkinson, the models, which do not require strictly rational behavior, are shown to describe the data well.

The conceptual device used to effect this extension is interesting in its own right. Statistical learning models have lent themselves particularly well to the study of partial-reinforcement situations in which the probability of reinforcement is contingent on the subject's choice. In a two-person game, this contingency is broadened to include the choice of both subjects, and it is not difficult to alter a model which applies in the one case so that it will apply in the other as well.

Additional References: Cervin and Henderson (1961); Suppes and Krasne (1961).

Concept Formation in Children

Suppes and Ginsberg (1962, 1963) set out in an ambitious program to study concept formation in children. In their 1962 study, for example, they applied the pattern model in an experiment in which children were to learn to discriminate between the numbers four and five presented successively and expressed in binary notation by means of a variety of symbolic expressions. (For example, the symbolic expressions $a\Delta\Delta$ and $a\Delta a$, respectively, represent in binary notation what in common notation are represented by the numerals 4 and 5.) They found that children may learn to discriminate between such expressions by identifying each one separately in terms of the particular combination of symbols in it rather than by assigning it to a class, the members of which share a common property. They also found that, different from adults in similar situations, children seem to learn faster—by showing appropriately different values of the parameter, which represents the effectiveness of each reinforcement on learning—when they are allowed to correct their mistakes than when they are only shown them.

Response Magnitude

Statistical learning theorists have studied sequences of zeros and ones (representing correct and incorrect responses) so exhaustively and exhaustingly that the use of other numbers to describe experimental results, as well as the use of other models to describe these numbers, has received much less attention than it deserves. In the period under review, the most important attempts to remedy this situation were made by Suppes and Frankmann (1961), Suppes and Zinnes (1961), and Anderson (1961).

The purpose of these studies was not to describe response latencies (descriptions of which have on rare enough occasions been attempted), but to describe response magnitudes. In the experiments by Suppes and Frankmann, for example, subjects were required to indicate where on the circumference of a circle a spot of light would subsequently fall, and in those experiments which Anderson's models are intended to describe, subjects were required to guess lengths of lines.

One of the results of the experiments by Suppes and Frankmann is of special interest. In these experiments, it was arranged that on every trial a light would fall on a point in an arc of fixed length in such a way that the probability that it would fall on any particular point in the arc was proportional to the height above that point of a triangle which could be constructed over the arc as its base. The shape of the histogram describing the asymptotic predictions of the subjects, however, was not triangular. This result is contrary to the matching rule which requires that reinforcement and response probabilities be asymptotically equal.

Reaction Time and Other Topics

LaBerge (1962) presented a model based on stimulus sampling ideas which is intended to account for results of reaction-time, simple-discrimination, and signal-detection experiments. The basic ideas of this model are that (a) on any experimental trial there is a constant proportion of stimuli which, when a sufficient number of them have been sampled, will lead to a recordable or correct response, (b) each stimulus is sampled randomly and one at a time from the entire population of stimuli present on the trial, and (c) each successive sampling requires the same amount of time. LaBerge showed the model to be consistent with both reaction-time and signal-detection data reported elsewhere in the literature.

Additional References: Atkinson (1961, 1963); Carterette (1961); Luce (1963); Shaffer (1962).

Other Models

McGuire (1961) presented a three-process, response-strength model to describe the data of paired-associate learning experiments. The three processes involve the development of connections between (a) external stimuli and internal responses, (b) internal stimuli (stimulus properties of the internal responses) and external responses, and (c) external stimulus-response pairs. As indicated by Bower (1962), a model such as this one of McGuire's may be applied to paired-associate learning experiments where not only the connections between the stimulus and response elements but also the response elements themselves are to be learned. In the experiments in which Bower's model has shown great success, the responses were not, like the stimuli, nonsense syllables, but were numbers, obviously already quite familiar to the subjects.

Additional References: Feigenbaum and Simon (1961, 1962); Fey (1961); Kanal (1962a, b); Lamperti and Suppes (1960); Levine (1963); Murdock (1960); Pubols (1962).

Prospectus

The immediate future should embrace not only more tests, alterations, and extensions of extant models but also the creation of new ones. The pattern and strategy models represent the latest signposts along a new and adventuresome road. The pattern model, which has given rise to so many statistics and which has been applied with such success and utility not only in the field of learning but also in other fields, stands as the greatest single achievement of the period under review. Plenty of room for extension and improvement remains, however, and can be seen from a consideration of the following facts: (a) the limited applicability of current models, which describe only behavior involving a choice between two or more responses; (b) the unreasonable assumption, relaxed only in a model which LaBerge introduced in 1959, that stimuli are conditioned exclusively to responses in which the experimenter is interested (*see* LaBerge, Greeno, and Peterson, 1962); (c) the absence of adequately testable models which describe behavior wherein there are individual differences; (d) the lack of knowledge of the sampling distributions of almost all the statistics derived from extant models—knowledge which is necessary for tests of goodness of fit; and (e) the treatment of learning in such a mechanical fashion that the greatest success of existing models has been and can only be expected to continue to be in situations so devoid of emotional involvement and intellectual challenge that all they require of adult human subjects is robot behavior.

Although the field is open and the soil has hardly been broken, there is already in the air a humming that betokens the building of models which, compared with current models, will be not only statistically stronger but also psychologically sounder and more widely applicable.

Bibliography

ANDERSON, NORMAN H. "Two Learning Models for Responses Measured on a Continuous Scale." *Psychometrika* 26: 391-403; December 1961.

ATKINSON, RICHARD C. "A Generalization of Stimulus Sampling Theory." *Psychometrika* 26: 281-90; September 1961.

BOWER, GORDON H. "Application of a Model to Paired-Associate Learning." *Psychometrika* 26: 255-80; September 1961.

BOWER, GORDON H. "An Association Model for Response and Training Variables in Paired-Associate Learning." *Psychological Review* 69: 34-53; January 1962.

ESTES, W. K. "Learning Theory and the New 'Mental Chemistry.'" *Psychological Review* 67: 207-23; July 1960.

ESTES, W. K. "New Developments in Statistical Behavior Theory: Differential Tests of Axioms for Associative Learning." *Psychometrika* 26: 73-84; March 1961.

ESTES, W. K.; HOPKINS, B. L.; and CROTHERS, E. J. "All-or-None and Conservation Effects in the Learning and Retention of Paired Associates." *Journal of Experimental Psychology* 60: 329-39; December 1960.

FELDMAN, JULIAN, and NEWELL, ALLEN. "A Note on a Class of Probability Matching Models." *Psychometrika* 26: 333-37; September 1961.

FLOOD, MERRILL M. "Stochastic Learning Theory Applied to Choice Experiments with Rats, Dogs, and Men." *Behavioral Science* 7: 289-314; July 1962.

JONES, JOAN E. "All-or-None Versus Incremental Learning." *Psychological Review* 69: 156-60; March 1962.

LABERGE, DAVID. "A Recruitment Theory of Simple Behavior." *Psychometrika* 27: 375-96; December 1962.

LABERGE, DAVID; GREENO, JAMES A.; and PETERSON, O. F. "Nonreinforcement and Neutralization of Stimuli." *Journal of Experimental Psychology* 63: 207-13; February 1962.

LEVINE, MARVIN. "Cue Neutralization: The Effects of Random Reinforcements upon Discrimination Learning." *Journal of Experimental Psychology* 63: 438-43; May 1962.

MCGUIRE, WILLIAM J. "A Multiprocess Model for Paired-Associate Learning." *Journal of Experimental Psychology* 62: 335-47; October 1961.

RESTLE, FRANK. "Statistical Methods for a Theory of Cue Learning." *Psychometrika* 26: 291-306; September 1961.

RESTLE, FRANK. "The Selection of Strategies in Cue Learning." *Psychological Review* 69: 329-43; July 1962.

SUPPES, PATRICK, and ATKINSON, RICHARD C. *Markov Learning Models for Multiperson Interactions.* Stanford, Calif.: Stanford University Press, 1960. 296 pp.

SUPPES, PATRICK, and FRANKMANN, RAYMOND W. "Test of Stimulus Sampling Theory for a Continuum of Responses with Unimodal Noncontingent Determinate Reinforcement." *Journal of Experimental Psychology* 61: 122-32; February 1961.

SUPPES, PATRICK, and GINSBERG, ROSE. "Application of a Stimulus Sampling Model to Children's Concept Formation with and Without Overt Correction Responses." *Journal of Experimental Psychology* 63: 330-36; April 1962.

SUPPES, PATRICK, and GINSBERG, ROSE. "A Fundamental Property of All-or-None Models, Binomial Distribution of Responses Prior to Conditioning, with Application to Concept Formation in Children." *Psychological Review* 70: 139-61; March 1963.

SUPPES, PATRICK, and SCHLAG-REY, MADELEINE. "Test of Some Learning Models for Double Contingent Reinforcement." *Psychological Reports* 10: 259-68; February 1962.

SUPPES, PATRICK, and ZINNES, JOSEPH L. "Stochastic Learning Theories for a Response Continuum with Non-Determinate Reinforcement." *Psychometrika* 26: 373-90; December 1961.

TRABASSO, THOMAS R. "Additivity of Cues in Discrimination Learning of Letter Patterns." *Journal of Experimental Psychology* 60: 83-88; August 1960.

TRABASSO, THOMAS R. "Stimulus Emphasis and All-or-None Learning in Concept Identification." *Journal of Experimental Psychology* 65: 398-406; April 1963.

Additional References

ANDERSON, NORMAN H., and WHALEN, RICHARD E. "Likelihood Judgments and Sequential Effects in a Two-Choice Learning Situation." *Journal of Experimental Psychology* 60: 111-20; August 1960.

ATKINSON, RICHARD C. "The Observing Response in Discrimination Learning." *Journal of Experimental Psychology* 62: 253-62; September 1961.

ATKINSON, RICHARD C. "A Variable Sensitivity Theory of Signal Detection." *Psychological Review* 70: 91-106; January 1963.

BOURNE, LYLE E., JR. "Long-Term Effects of Misinformative Feedback upon Concept Identification." *Journal of Experimental Psychology* 65: 139-47; February 1963.

BOURNE, LYLE E., JR., and BUNDERSON, C. VICTOR. "Effects of Delay of Informative Feedback and Length of Postfeedback Interval on Concept Identification." *Journal of Experimental Psychology* 65: 1-5; January 1963.

CARTERETTE, TERESA S. "An Application of Stimulus Sampling Theory to Summated Generalization." *Journal of Experimental Psychology* 62: 448-55; November 1961.

CERVIN, V. B., and HENDERSON, G. P. "Statistical Theory of Persuasion." *Psychological Review* 68: 157-66; May 1961.

CHAPMAN, JEAN P. "The Spacing of Sequentially Dependent Trials in Probability Learning." *Journal of Experimental Psychology* 62: 545-51; December 1961.

DERKS, PETER L. "The Generality of the 'Conditioning Axiom' in Human Binary Prediction." *Journal of Experimental Psychology* 63: 538-45; June 1962.

DERKS, PETER L. "Effect of Run Length on the 'Gambler's Fallacy.'" *Journal of Experimental Psychology* 65: 213-14; February 1963.

EDWARDS, WARD. "Probability Learning in 1000 Trials." *Journal of Experimental Psychology* 62: 385-94; October 1961.

ESTES, W. K., and HOPKINS, B. L. "Acquisition and Transfer in Pattern-vs.-Component Discrimination Learning." *Journal of Experimental Psychology* 61: 322-28; April 1961.

FEIGENBAUM, EDWARD A., and SIMON, HERBERT A. "Comment: The Distinctiveness of Stimuli." *Psychological Review* 68: 285-88; July 1961.

FEIGENBAUM, EDWARD A., and SIMON, HERBERT A. "A Theory of the Serial Position Effect." *British Journal of Psychology* 53: 307-20; August 1962.

FEY, CURT F. "An Investigation of Some Mathematical Models for Learning." *Journal of Experimental Psychology* 61: 455-61; June 1961.

GREENO, JAMES G. "Effects of Nonreinforced Trials in Two-Choice Learning with Noncontingent Reinforcement." *Journal of Experimental Psychology* 64: 373-79; October 1962.

HICKSON, ROBERT H. "Response Probability in a Two-Choice Learning Situation with Varying Probability of Reinforcement." *Journal of Experimental Psychology* 62: 138-44; August 1961.

KANAL, LAVEEN. "The Asymptotic Distribution for the Two-Absorbing Barrier Beta Model." *Psychometrika* 27: 105-109; March 1962. (a)

KANAL, LAVEEN. "A Functional Equation Analysis of Two Learning Models." *Psychometrika* 27: 89-104; March 1962. (b)

LAMPERTI, JOHN, and SUPPES, PATRICK. "Some Asymptotic Properties of Luce's Beta Learning Model." *Psychometrika* 25: 233-41; September 1960.

LEVINE, MARVIN. "Mediating Processes in Humans at the Outset of Discrimination Learning." *Psychological Review* 70: 254-76; May 1963.

LUCE, R. DUNCAN. "A Threshold Theory for Simple Detection Experiments." *Psychological Review* 70: 61-79; January 1963.

MCCRACKEN, J.; OSTERHOUT, C.; and VOSS, JAMES F. "Effects of Instructions in Probability Learning." *Journal of Experimental Psychology* 64: 267-71; September 1962.

MURDOCK, BENNET B., JR. "The Immediate Retention of Unrelated Words." *Journal of Experimental Psychology* 60: 222-34; October 1960.

NIES, RICHARD C. "Effects of Probable Outcome Information on Two-Choice Learning." *Journal of Experimental Psychology* 64: 430-33; November 1962.

PETERSON, LLOYD R., and OTHERS. "Recency and Frequency in Paired-Associate Learning." *Journal of Experimental Psychology* 63: 396-403; April 1962.

PISHKIN, VLADIMIR. "Effects of Probability of Misinformation and Number of Irrelevant Dimensions upon Concept Identification." *Journal of Experimental Psychology* 59: 371-78; June 1960.

PROKASY, WILLIAM F., JR. "Nonrandom Stimulus Sampling in Statistical Learning Theory." *Psychological Review* 68: 219-24; May 1961.

PUBOLS, BENJAMIN H., JR. "An Application of Levine's Model for Hypothesis Behavior to Serial Reversal Learning." *Psychological Review* 69: 241-45; May 1962.

REYNOLDS, JAMES H. "Repetition Factors in One-Trial Learning of Paired Associates." *Journal of Experimental Psychology* 65: 352-67; April 1963.

RICHARDS, JAMES M., JR. "The Cue Additivity Principle in a Restricted Social Interaction Situation." *Journal of Experimental Psychology* 63: 452-57; May 1962.

SHAFFER, JULIET POPPER. "Discrimination and Mediated Generalization in Probability Learning." *Journal of Experimental Psychology* 64: 593-99; December 1962.

SHAFFER, JULIET POPPER. "Effect of Different Stimulus Frequencies on Discrimination Learning with Probabilistic Reinforcement." *Journal of Experimental Psychology* 65: 265-69; March 1963.

SIMON, HERBERT A. "A Note on Mathematical Models for Learning." *Psychometrika* 27: 417-18; December 1962.

SUPPES, PATRICK, and KRASNE, FRANKLIN. "Application of Stimulus Sampling Theory to Situations Involving Social Pressure." *Psychological Review* 68: 46-59; January 1961.

WAUGH, NANCY C., and SMITH, J. E. KEITH. "A Stochastic Model for Free Recall." *Psychometrika* 27: 141-54; June 1962.

WILLIAMS, JOANNA P. "A Test of the All-or-None Hypothesis for Verbal Learning." *Journal of Experimental Psychology* 64: 158-65; August 1962.

WITTE, ROBERT S. "Conditional Response Probability in a T Maze." *Journal of Experimental Psychology* 62: 439-47; November 1961.

CHAPTER IX

Decision Theory and Utility Management

ROBERT W. EARL

SINCE Girshick presented his elementary survey of statistical decision theory in the December 1954 issue of the REVIEW, a tremendous amount of research has been published by behavioral scientists in journals that fall outside the focus of most workers in educational research. The new mathematical psychologists who have made important methodological contributions to the understanding of problems in decision theory and utility measurement have not been very much concerned with problems of immediate or practical interest to the educational psychologist, the curriculum specialist, the counselor, or the school administrator. Once many of the methodological problems have been resolved and a substantial amount of experimental work has been done to ascertain the suitability of the somewhat esoteric models in decision theory to the understanding and to the potential solution of certain educational problems, there will probably be little use of decision theory by research specialists in education. In the writer's estimation, at least 10 or 15 years will elapse before any marked application of decision theory will be made to the types of educational problems faced by most educational researchers.

In view of the substantial overlap of the field of decision theory with (a) the problems of statistical inference, considered by Clark in Chapter I and by Michael in Chapter II; (b) the problems in psychophysics, signal detection, and information theory as well as general developments in scaling and measurement theory, presented by Schiffman and Messick in Chapter VII; (c) statistical learning models, detailed by Weitzman in Chapter VIII; and (d) computer technology, as described by Baker in Chapter X, the writer decided with somewhat mixed feelings that his presentation should be drastically curtailed in length and in scope relative to what he had initially planned to undertake. Nevertheless, a detailed survey was made of pertinent literature since about 1950. For the most part, references since 1954 have been included. Probably the most important contribution of this chapter will be to furnish the reader with a bibliography of slightly more than 100 carefully selected references which furnish what may be considered an appropriate base for the acquisition of an understanding of decision theory and utility management in the behavioral sciences. These references were selected on the basis of extensive consultation with several mathematical psychologists who have specialized in the study of decision theory.

General References

The literature of decision theory varies considerably in the mathematical sophistication it demands of its readers. The interested reader can study the theory of decision in considerable depth by consulting the following references, which, although differing in levels of mathematical maturity required for their comprehension, tend to be general in treatment: Adams (1960), Bakan (1953), Bross (1953), Chernoff and Moses (1959), Edwards (1954, 1961, 1962a, b, c), Edwards, Lindman, and Savage (1963), Raiffa and Schlaifer (1961), Savage (1954), Schlaifer (1959), Von Neumann and Morgenstern (1947), and Wald (1947, 1950).

Some Basic Considerations

Recent books and articles on "statistical decision" theory contain a number of remarks at variance with the fiat and scripture that currently dictate the practice of statistical generalization in scientific research and industrial quality control. For example, Schlaifer (1959, p. 629), in a moment of restraint, stated that traditional hypothesis testing, which in his opinion is worse than useless, leads to absurd and dangerous conclusions. Similarly, Savage (1962, p. 181) stated that it is not unusual for a t value of three with many degrees of freedom to *support* the null hypothesis. Raiffa and Schlaifer (1961, p. 41) stated that an experimenter may terminate an experiment for "whatever reason," provided that the manner of the data analysis is Bayesian. For stopping an experiment, Edwards, Lindman, and Savage (1963) listed such possible "reasons whatever" as being tired, answering the telephone, running out of money, feeling one has sufficient data to prove a point, but not necessarily being chased away by lions.

In his most recent joust with folklore, Binder (1963) took exception to Grant's (1962) argument that rejection of the null hypothesis can support an empirical hypothesis while acceptance of the null hypothesis cannot. Grant's point of view possibly derived from the parallel suggestion of Francis Bacon that scientific induction is an asymmetrical process under which enumeration of confirming instances is "exposed to danger from one contradictory instance" while reliance on negative instances to eliminate false hypotheses is not.

As the first psychologist since Lewin and his associates (Cartwright, 1941; Festinger, 1942, 1943) to become inseparably identified with decision theory, Edwards (1961) reacted to recent developments in mathematical statistics by issuing a call for someone to lead psychologists out of the wilderness of t-tests and F-tests. Failing to elicit a response to his call, he answered the call himself (Edwards, Lindman, and Savage, 1963).

Although the authors of the preceding statements refer to hypothesis testing and traditional tests of significance, they are actually arguing within the slightly broader context of ethics, aesthetics, logic, metaphysics, epistemology, ampliative induction, inconclusive argument, statistical inferences, decision theory, and behavior theory, whether of man, mouse, or machine, to give but a small sample of the topics upon which their theories bear. Nominally, the articles cited in the bibliography in this chapter are largely directed toward the study of (a) optimal gambling behavior; (b) optimal economic decision making; (c) optimal human, animal, or machine decision making, regardless of context; (d) common day-to-day, infrequently optimal decision-making behavior of human beings or animals; (e) human data processing; and (f) classical versus Bayesian statistical inference. By contrast with the breadth of the subject matter it entails, the theory is quite simply called decision theory.

Central Concepts, Questions, and Difficulties

The origins of decision theory can be traced to the unsuccessful attempts of early Renaissance mathematicians to answer questions raised in connection with commercial insurance and gambling. The first successful conceptualization of the problem of decision making came later, primarily with the work of Laplace. The major concepts developed during the eighteenth century were those of probability and expectation. Attempts to define probability in a way satisfactory to everyone are still in progress and are particularly pertinent to decision theory.

A concept such as probability, which is ordinarily used both as a term within an abstract system and as a synonym for the relative frequency of occurrence of certain definitive events of the real world, is a source of confusion even for experts. To add to the confusion, psychologists such as Festinger (1942), Atkinson (1957), and Edwards (1962b) have added behavioral definitions of probability to those mentioned above. Even Laplace and later mathematicians, statisticians, and economists had talked about measures of credibility or credence in lieu of the objective frequency definition whenever the objective approach created "logical" difficulties.

Probability

In the mathematical formulation, probability is treated as an undefined term that derives its meaning indirectly from the way that the term is manipulated in accordance with the rules of the mathematical system. Parenthetically, the concept of randomness is treated similarly in the system. Coombs, Raiffa, and Thrall (1954) presented a very clear and understandable discussion of the distinction between probability as an abstract mathematical entity and probability defined with reference to experimental events. Probability, as the undefined component of a mathematical system, is no problem in the present context. The problem faced by decision theorists is choosing an experimental definition of probability that will do what they want it to do and then sticking to it.

The problem of the reader is even more difficult than that of the theorist. The reader must discover the theorist's definition, if this is possible, and then must make sure that the theorist adheres to it. Since the reader is demanding the impossible, he must catch the theorist each time he shifts to a new definition and finally must determine how many different definitions were used and precisely what they were. Experience suggests that the burden is on the reader, for the theorists, in this field at least, have more important problems than to satisfy the reader, as one may see from the next few paragraphs.

Mathematical Expectation

Mathematical expectation is the second major concept in decision theory. Mathematical expectation, or expected value, is simply a synonym for the mean, or average. For example, the scores of three students on a final examination are 2, 5, and 8; the expected score is 5. This profound concept, and it is exactly that, is symbolized $E(x) = 5$, where x represents a score. The importance of this concept to decision theory is that a decision maker faced with the necessity of choosing between two or more alternatives will choose the alternative associated with the maximum (or minimum) expected value. Given a choice between teaching a class with $E(x) = 5$, $E(x) = 6$, $E(x) = 7$, a decision maker would choose the latter if he valued the scores in proportion to their magnitude. If he valued low scores more than high scores, however, the decision maker would choose $E(x) = 5$. This example merely illustrates the importance of the concept of expectation to decision theory; it does not justify the decision.

With no special provocation, one could argue that the environment that gave rise to decision theory has determined, or more precisely has restricted, the selection of the variables upon which the theory is based. The probabilities obviously derive from games of chance; the scores, or x values, are simply amounts of money (e.g., dollars, rupees); and the expected values, $E(x)$, are simply the amounts of money one is likely to win or lose per play in the long run. The decision criterion is to bet on the alternative that, on the average, will pay the most (or will lose the least, if winning is impossible and one is forced to play). The rebuttal to this argument is that in order to offer more than "good" advice to gamblers (whatever form the gambling activity takes) and to apply this conceptualization to areas of scientific interest, the decision theorist has had to modify his theory in order to come to terms with reality. Thus, he has freed the theory from the original restrictions. For example, theory was applied to the prediction of human decision making in gambling situations as opposed to the usual normative application of telling the gambler how he should behave. Since the results were disastrous, the theory was modified in the simplest way possible. The axioms were not touched, but the identifications between the abstract concepts and the real world events were modified, with the result that to date (a) probability refers to the subject's expressed impression of the chances that a given outcome, such as heads, will occur; (b) the value x now indicates the subject's expressed preference for the outcome relative to the others; and (c) the expected value is simply a preference for an act of choice determined by the same formula by which one would compute the mean score of the students in a class.

Conclusion

Three comments should be made. First, in its revised form, decision theory is still normative in that it predicts that a decision maker will behave optimally in terms of his subjective interpretation of the pleasures and chances in his environment. Second, the theory is applied with complete equanimity—on those rare occasions when it is applied—to any kind of decision maker worth imagining (the unworthy imaginables are excluded); all living creatures and most machines are included. Third, the theory is a creation of mathematicians and of mathematically oriented statisticians and economists. The sole preoccupation is the development of consistent axiom systems and theorems that can be deduced from them. Such a system, the terms of which are identified with experimental events, is what one calls, in a current context, a theory; and a theory can be tested against the phenomena it is supposed to describe. If the theory fails often enough, the theorist can modify it by changing the axiomatic basis or the identifications, or he can replace the model.

Now the key point to be made is that the tests of decision theory against data are virtually nonexistent. Most of the experiments—what few there are—test implications of the various theories of measurement of the subjective values (preferabilities) of the x's. These preferabilities, which are called *utilities* in decision theory, basically represent nothing more than a rank ordering of the x values of the outcomes and a simple (rank) or partial ordering of all pairs of outcomes (on the basis of the subject's estimates of the subjective magnitude of the differences between the x values). The "tests" of the theory have taken the form of esoteric anecdotes dreamed up in the armchair by the theorists. (A case in point is the famous St. Petersburg paradox, which may be left to the reader's curiosity.) These anecdotes have influenced the choice of identifications almost exclusively. On the other hand, the variations in the axiom systems have come about primarily by the need of new decision theorists to make an original contribution to theory.

Bibliography

ADAMS, ERNEST W. "Survey of Bernoullian Utility Theory." *Mathematical Thinking in the Measurement of Behavior.* (Edited by Herbert Solomon.) Glencoe, Ill.: Free Press, 1960. Part 2, pp. 151-268.

ATKINSON, JOHN W. "Motivational Determinants of Risk-Taking Behavior." *Psychological Review* 64: 359-72; November 1957.

BAKAN, DAVID. "Learning and the Principle of Inverse Probability." *Psychological Review* 60: 360-70; November 1953.

BINDER, ARNOLD. "Further Considerations on Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 70: 107-15; January 1963.

BROSS, IRWIN D. J. *Design for Decision.* New York: Macmillian Co., 1953. 276 pp.

CARTWRIGHT, DORWIN. "Decision-Time in Relation to the Differentiation of the Phenomenal Field." *Psychological Review* 48: 425-42; September 1941.

CHERNOFF, HERMAN, and MOSES, LINCOLN E. *Elementary Decision Theory.* New York: John Wiley & Sons, 1959. 364 pp.

COOMBS, C. H.; RAIFFA, H.; and THRALL, R. M. "Some Views on Mathematical Models and Measurement Theory." *Decision Processes.* (Edited by R. M. Thrall, C. H. Coombs, and R. L. Davis.) New York: John Wiley & Sons, 1954. Chapter 2, pp. 19-37.

EDWARDS, WARD. "The Theory of Decision Making." *Psychological Bulletin* 51: 380-417; July 1954.

EDWARDS, WARD. "Behavioral Decision Theory." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1961. Vol. 12, pp. 473-98.

EDWARDS, WARD. "Dynamic Decision Theory and Probabilistic Information Processing." *Human Factors* 4: 59-73; April 1962. (a)

EDWARDS, WARD. "Subjective Probabilities Inferred from Decisions." *Psychological Review* 69: 109-35; March 1962. (b)

EDWARDS, WARD. "Utility, Subjective Probability, Their Interaction, and Variance Preferences." *Journal of Conflict Resolution* 6: 42-51; March 1962. (c)

EDWARDS, WARD; LINDMAN, HAROLD; and SAVAGE, LEONARD J. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70: 193-242; May 1963.

FESTINGER, LEON. "A Theoretical Interpretation of Shifts in Level of Aspiration." *Psychological Review* 49: 235-50; May 1942.

FESTINGER, LEON. "Studies in Decision: I. Decision-Time, Relative Frequency of Judgment and Subjective Confidence as Related to Physical Stimulus Difference." *Journal of Experimental Psychology* 32: 291-306; April 1943.

GRANT, DAVID A. "Testing the Null Hypothesis and the Strategy and Tactics of Investigating Theoretical Models." *Psychological Review* 69: 54-61; January 1962.

RAIFFA, HOWARD, and SCHLAIFER, ROBERT. *Applied Statistical Decision Theory.* Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1961. 356 pp.

SAVAGE, LEONARD J. *The Foundations of Statistics.* New York: John Wiley & Sons, 1954. 294 pp.

SAVAGE, LEONARD J. "Bayesian Statistics." *Recent Developments in Information and Decision Processes.* (Edited by Robert E. Machel and Paul Gray.) New York: Macmillan Co., 1962. pp. 161-94.

SCHLAIFER, ROBERT. *Probability and Statistics for Business Decisions.* New York: McGraw-Hill Book Co., 1959. 732 pp.

VON NEUMANN, JOHN, and MORGENSTERN, OSKAR. *Theory of Games and Economic Behavior.* Second edition. Princeton, N.J.: Princeton University Press, 1947. 641 pp.

WALD, ABRAHAM. *Sequential Analysis.* New York: John Wiley & Sons, 1947. 212 pp.

WALD, ABRAHAM. *Statistical Decision Functions.* New York: John Wiley & Sons, 1950. 179 pp.

Additional References

ATKINSON, JOHN W., and OTHERS. "The Achievement Motive, Goal Setting, and Probability Preferences." *Journal of Abnormal and Social Psychology* 60: 27-36; January 1960.

ATKINSON, RICHARD C. "A Variable Sensitivity Theory of Signal Detection." *Psychological Review* 70: 91-106; January 1963.

BECKER, GORDON M. "Sequential Decision Making: Wald's Model and Estimates of Parameters." *Journal of Experimental Psychology* 55: 628-36; June 1958.

BECKER, GORDON M. "Decision Making: Objective Measures of Subjective Probability and Utility." *Psychological Review* 69: 136-48; March 1962.

BERLYNE, D. E. "Motivational Problems Raised by Exploratory and Epistemic Behavior." *Psychology: A Study of Science. The Process Areas, the Person, and Some Applied Fields: Their Place in Psychology and in Science.* (Edited by Sigmund Koch.) New York: McGraw-Hill Book Co., 1963. Vol. 5, pp. 284-364.

BERRY, RICHARD N. "On Lotsof's Proposed Relation Between Certainty and Expectancy." *American Journal of Psychology* 76: 333-34; June 1963.

BLACKWELL, H. RICHARD. "Neural Theories of Simple Visual Discriminations." *Journal of the Optical Society of America* 53: 129-60; January 1963.

BUSH, ROBERT R., and MOSTELLER, FREDERICK. *Stochastic Models for Learning.* New York: John Wiley & Sons, 1955. 365 pp.

BUSH, ROBERT R.; GALANTER, EUGENE; and LUCE, R. DUNCAN. "Characterization and Classification of Choice Experiments." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, Chapter 2, pp. 77-102.

COLUMBIA UNIVERSITY, STATISTICAL RESEARCH GROUP. "Sequential Analysis When the Result of a Single Observation Is a Classification as Good or Bad and When the Result of the Test Is a Decision Between Two Methods or Products (Double Dichotomies)." *Sequential Analysis of Statistical Data: Applications.* SRG Report 255. New York: Columbia University Press, 1945. Section 3 (revised), pp. 3.01-3.57.

COOMBS, C. H. "An Application of a Nonmetric Model for Multidimensional Analysis of Similarities." *Psychological Reports* 4: 511-18; September 1958.

COOMBS, C. H., and KOMORITA, S. S. "Measuring Utility of Money Through Decisions." *American Journal of Psychology* 71: 383-89; June 1958.

COWAN, THOMAS A. "Decision Theory in Law, Science, and Technology." *Science* 140: 1065-75; June 7, 1963.

DALE, H. C. A. "Strategies of Searching in Two Simple Systems." *American Journal of Psychology* 72: 539-46; December 1959.

DAVIDSON, DONALD, and SUPPES, PATRICK. *Decision Making: An Experimental Approach.* Stanford, Calif.: Stanford University Press, 1957. 121 pp.

DAVIS, ROBERT L. "Introduction to 'Decision Processes.'" *Decision Processes.* (Edited by R. M. Thrall, C. H. Coombs, and R. L. Davis.) New York: John Wiley & Sons, 1954. Chapter 1, pp. 1-18.

DEMBER, WILLIAM N. "Alternation Behavior." *Functions of Varied Experience.* (Compiled by Donald W. Fiske and Salvatore R. Maddi.) New York: Dorsey Press, 1961. Chapter 8, pp. 227-52.

DEMBER, WILLIAM N., and EARL, ROBERT W. "Analysis of Exploratory, Manipulatory, and Curiosity Behaviors." *Psychological Review* 64: 91-96; March 1957.

DETAMBEL, MARVIN H., and STOLUROW, LAWRENCE M. "Probability and Work as Determiners of Multichoice Behavior." *Journal of Experimental Psychology* 53: 73-81; February 1957.

EARL, R. W. *A Theory of Stimulus Selection.* Special Document SD61-132. Fullerton, Calif.: Hughes Aircraft Company, Ground Systems, 1961. 31 pp.

EDWARDS, WARD. "The Prediction of Decisions Among Bets." *Journal of Experimental Psychology* 50: 201-14; September 1955.

EDWARDS, WARD. "Measurement of Utility and Subjective Probability." *Psychological Scaling: Theory and Applications.* (Edited by Harold Gulliksen and Samuel Messick.) New York: John Wiley & Sons, 1960. Chapter 10, pp. 109-27.

EDWARDS, WARD. "Costs and Payoffs Are Instructions." *Psychological Review* 68: 275-84; July 1961. (a)

EDWARDS, WARD. "Probability Learning in 1000 Trials." *Journal of Experimental Psychology* 62: 385-94; October 1961. (b)

EGAN, JAMES P., and CLARKE, FRANK R. *Psychophysics and Signal Detection.* Final Technical Report, Contract No. AF 19(628)-266, Report No. ESD-TDR-62-305. Bloomington: Indiana University Press, 1962. 92 pp.

ESTES, W. K. "Of Models and Men." *American Psychologist* 12: 609-17; October 1957.

ESTES, W. K. "The Statistical Approach to Learning Theory." *Psychology: A Study of Science. General Systematic Formulations, Learning, and Special Processes.* (Edited by Sigmund Koch.) New York: McGraw-Hill Book Co., 1959. Vol. 2, pp. 380-491.

FISKE, DONALD W., and JONES, LYLE V. "Sequential Analysis in Psychological Research." *Psychological Bulletin* 51: 264-75; May 1954.

FLOOD, MERRILL M. "On Game-Learning Theory and Some Decision-Making Experiments." *Decision Processes.* (Edited by R. M. Thrall, C. H. Coombs, and R. L. Davis.) New York: John Wiley & Sons, 1954. Chapter 10, pp. 139-58.

GALANTER, E. H., and SMITH, W. A. S. "Some Experiments on a Simple Thought-Problem." *American Journal of Psychology* 71: 359-66; June 1958.

GARDNER, R. ALLEN. "Multiple-Choice Decision-Behavior." *American Journal of Psychology* 71: 710-17; December 1958.

GARNER, WENDELL R. *Uncertainty and Structure as Psychological Concepts.* New York: John Wiley & Sons, 1962. 369 pp.

GOODNOW, JACQUELINE JARRETT. "Determinants of Choice-Distribution in Two-Choice Situations." *American Journal of Psychology* 68: 106-16; March 1955.

GREEN, BERT F., JR.; WOLF, ALICE K.; and WHITE, BENJAMIN W. "The Detection of Statistically Defined Patterns in a Matrix of Dots." *American Journal of Psychology* 72: 503-20; December 1959.

GREEN, DAVID M. "Psychoacoustics and Detection Theory." *Journal of the Acoustical Society of America* 32: 1189-1203; October 1960.

GREEN, DAVID M. "Some Comments and a Correction of 'Psychoacoustics and Detection Theory.'" *Readings in Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, pp. 67-68.

HACK, MARTIN H. "Signal Detection in the Rat." *Science* 139: 758-60; February 22, 1963.

HAYES, SAMUEL P. "Some Psychological Problems of Economics." *Psychological Bulletin* 47: 289-330; July 1950.

HAYS, WILLIAM L. *Statistics for Psychologists.* New York: Holt, Rinehart and Winston, 1963. 719 pp.

HELLYER, SIDNEY, and STRAUGHAN, JAMES H. "Alternation as Function of Preliminary Training and Type of Deprivation." *Science* 133: 1422-23; May 5, 1961.

HOWARD, THOMAS C. "The Relation Between Psychological and Mathematical Probability." *American Journal of Psychology* 76: 335; June 1963.

IRWIN, FRANCIS W. "An Analysis of the Concepts of *Discrimination* and *Preference.*" *American Journal of Psychology* 71: 152-63; March 1958.

IRWIN, FRANCIS W., and SMITH, W. A. S. "Value, Cost and Information as Determiners of Decision." *Journal of Experimental Psychology* 54: 229-32; September 1957.

JONES, JOAN E. "All-or-None Versus Incremental Learning." *Psychological Review* 69: 156-60; March 1962.

KAPLAN, RICHARD J., and NEWMAN, J. ROBERT. *A Study in Probabilistic Information Processing (PIP).* Technical Memorandum TM-1150-000-00. Santa Monica, Calif.: System Development Corporation, 1963. 36 pp.

KERSHNER, R. B., and WILCOX, L. R. *The Anatomy of Mathematics.* New York: Ronald Press, 1950. 416 pp.

LAWRENCE, DOUGLAS H., reviewer. "Review of 'Modern Learning Theory' by William K. Estes and Others." *American Journal of Psychology* 68: 152-56; March 1955.

LEWIS, DONALD J. "Partial Reinforcement: A Selective Review of the Literature Since 1950." *Psychological Bulletin* 57: 1-28; January 1960.

LICKLIDER, J. C. R. "Three Auditory Theories." *Psychology: A Study of Science. Sensory, Perceptual, and Physiological Formulations.* (Edited by Sigmund Koch.) New York: McGraw-Hill Book Co., 1959. Vol. 1, pp. 41-144.

LITTLE, KENNETH B., and OTHERS. "A Further Test of a General Utility Theory Model for Probability Learning." *Journal of Experimental Psychology* 66: 107-108; July 1963.

LOTSOF, ERWIN J. "Expectancy for Success and Decision-Time." *American Journal of Psychology* 71: 416-19; June 1958.

LOTSOF, ERWIN J. "Expectancy for Success and Certainty of Response." *American Journal of Psychology* 72: 600-602; December 1959.

LUBIN, ARDIE. "Statistics." *Annual Review of Psychology.* (Edited by Paul R. Farnsworth, Olga McNemar, and Quinn McNemar.) Palo Alto, Calif.: Annual Reviews, 1962. Vol. 13, pp. 345-70.

LUCE, R. DUNCAN. *Individual Choice Behavior: A Theoretical Analysis.* New York: John Wiley & Sons, 1959. 153 pp. (a)

LUCE, R. DUNCAN. "A Probabilistic Theory of Utility and Its Relationship to Fechnerian Scaling." *Measurement: Definitions and Theories.* (Edited by C. West Churchman and Philburn Ratoosh.) New York: John Wiley & Sons, 1959. Chapter 7, pp. 144-59. (b)

LUCE, R. DUNCAN. "Detection and Recognition." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, Chapter 3, pp. 103-89. (a)

LUCE, R. DUNCAN. "On the Possible Psychophysical Laws." *Readings in Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, pp. 69-83. (b)

LUCE, R. DUNCAN. "A Threshold Theory for Simple Detection Experiments." *Psychological Review* 70: 61-79; January 1963. (c)

LUCE, R. DUNCAN, and GALANTER, EUGENE. "Discrimination." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, Chapter 4, pp. 191-243.

LUCE, R. DUNCAN, and RAIFFA, HOWARD. *Games and Decisions: Introduction and Critical Survey.* New York: John Wiley & Sons, 1957. 509 pp.

LUCE, R. DUNCAN; BUSH, ROBERT R.; and GALANTER, EUGENE, editors. *Handbook of Mathematical Psychology.* New York: John Wiley & Sons, 1963. Vol. 1, 491 pp. (a)

LUCE, R. DUNCAN; BUSH, ROBERT R.; and GALANTER, EUGENE, editors. *Readings in Mathematical Psychology.* New York: John Wiley & Sons, 1963. Vol. 1, 535 pp. (b)

MCGLOTHLIN, WILLIAM H. "Stability of Choices Among Uncertain Alternatives." *American Journal of Psychology* 69: 604-15; December 1956.

MILNOR, JOHN. "Games Against Nature." *Decision Processes.* (Edited by R. M. Thrall, C. H. Coombs, and R. L. Davis.) New York: John Wiley & Sons, 1954. Chapter 4, pp. 49-59.

MORSE, ELIZABETH B., and RUNQUIST, WILLARD N. "Probability-Matching with an Unscheduled Random Sequence." *American Journal of Psychology* 73: 603-607; December 1960.

PARZEN, EMANUEL. *Modern Probability Theory and Its Applications.* New York: John Wiley & Sons, 1960. 464 pp.

PETERSON, W. W.; BIRDSALL, T. G.; and FOX, W. C. "The Theory of Signal Detectability." *Readings in Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, pp. 167-211.

RAIFFA, HOWARD. "Bayesian Decision Theory." *Recent Developments in Information and Decision Processes.* (Edited by Robert E. Machol and Paul Gray.) New York: Macmillan Co., 1962. pp. 92-101.

SCHUSTERMAN, RONALD J. "The Use of Strategies in Two-Choice Behavior of Children and Chimpanzees." *Journal of Comparative and Physiological Psychology* 56: 96-100; February 1963.

SCOTT, DANA, and SUPPES, PATRICK. "Foundational Aspects of Theories of Measurement." *Readings in Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, pp. 212-27.

SIEGEL, SIDNEY. "Theoretical Models of Choice and Strategy Behavior: Stable State Behavior in the Two-Choice Uncertain Outcome Situation." *Psychometrika* 24: 303-16; December 1959.

SIEGEL, SIDNEY. "Individual Decision-Making Under Risk: A Model for Finite Choice Structures." *Psychological Scaling: Theory and Applications.* (Edited by Harold Gulliksen and Samuel Messick.) New York: John Wiley & Sons, 1960. Chapter 11, pp. 129-43.

SIEGEL, SIDNEY, and FOURAKER, LAWRENCE E. *Bargaining and Group Decision Making.* New York: McGraw-Hill Book Co., 1960. 132 pp.

STARBUCK, WILLIAM H. "Level of Aspiration." *Psychological Review* 70: 51-60; January 1963.

STEVENS, S. S. "Measurement, Psychophysics, and Utility." *Measurement: Definitions and Theories.* (Edited by C. West Churchman and Philburn Ratoosh.) New York: John Wiley & Sons, 1959. Chapter 2, pp. 18-63.

SUPPES, PATRICK. *Concept Formation and Bayesian Decisions.* Technical Memorandum TM-1262. Santa Monica, Calif.: System Development Corporation, 1963. 31 pp.

SUPPES, PATRICK, and ZINNES, JOSEPH L. "Basic Measurement Theory." *Handbook of Mathematical Psychology.* (Edited by R. Duncan Luce, Robert R. Bush, and Eugene Galanter.) New York: John Wiley & Sons, 1963. Vol. 1, Chapter 1, pp. 1-76.

SWETS, JOHN A. "Detection Theory and Psychophysics: A Review." *Psychometrika* 26: 49-63; March 1961. (a)

SWETS, JOHN A. "Is There a Sensory Threshold?" *Science* 134: 168-77; July 21, 1961. (b)

SWETS, JOHN A. "Information Retrieval Systems." *Science* 141: 245-50; July 19, 1963.

SWETS, JOHN A., and GREEN, DAVID M. "Sequential Observations by Human Observers of Signals in Noise." *Information Theory.* (Edited by Colin Cherry.) London: Butterworth Scientific Publications, 1961. Chapter 15, pp. 177-95.

SWETS, JOHN A., and SEWALL, SUSAN T. "Stimulus vs. Response Uncertainty in Recognition." *Journal of the Acoustical Society of America* 33: 1586-92; November 1961.

SWETS, JOHN A.; TANNER, WILSON P., JR.; and BIRDSALL, THEODORE G. "Decision Processes in Perception." *Psychological Review* 68: 301-40; September 1961.

THRALL, R. M.; COOMBS, C. H.; and DAVIS, R. L., editors. *Decision Processes.* New York: John Wiley & Sons, 1954. 332 pp.

TOLMAN, EDWARD C. *Purposive Behavior in Animals and Men.* Berkeley: University of California Press, 1949. 463 pp.

TOLMAN, EDWARD C. *Collected Papers in Psychology.* Berkeley: University of California Press, 1951. 269 pp.

TOMBAUGH, TOM N., and MARX, MELVIN H. "The Effects of Increasing and Decreasing Orders of Monetary Incentive on Resistance to Extinction in Children." *Psychological Record* 13: 187-90; April 1963.

VERPLANCK, WILLIAM S. *A Glossary of Some Terms Used in the Objective Science of Behavior.* Psychological Review, Vol. 64, No. 6, Part 2, Supplement. Washington, D.C.: American Psychological Association, November 1957. 42 pp.

WALKER, EDWARD L. "Action Decrement and Its Relation to Learning." *Psychological Review* 65: 129-42; May 1958.

WALKER, E. L., and OTHERS. "Choice Alternation: I. Stimulus vs. Place vs. Response." *Journal of Comparative and Physiological Psychology* 48: 19-23; February 1955. (a)

WALKER, E. L., and OTHERS. "Choice Alternation: III. Response Intensity vs. Response Discriminability." *Journal of Comparative and Physiological Psychology* 48: 80-85; February 1955. (b)

WILLIAMS, J. D. *The Compleat Strategyst.* New York: McGraw-Hill Book Co., 1954. 234 pp.

WYLIE, CLARENCE R., JR. "Paradox." *Scientific Monthly* 67: 63; July 1948.

CHAPTER X

Use of Computers in Educational Research

FRANK B. BAKER

SINCE Martin and Hall wrote their chapter concerning data processing for the December 1960 REVIEW, the computer field has been kaleidoscopic, with new developments appearing and older trends growing to new levels. Perhaps the most marked change has been in the equipment field. In the previous REVIEW period, the basic computer was the IBM 650, although a few major universities either had built their own large-scale computers or had acquired a large-scale machine returned to the manufacturer by the aerospace industry. A significant change in this pattern has arisen, primarily as a result of the convergence of four factors: (a) availability of the transistorized computer, (b) change in the fiscal point of view within university administrations, (c) substantial educational discounts (up to 60 percent) from the manufacturers, and (d) significant financial support from the National Science Foundation. Because of these four factors, most major universities have installed the newest and most powerful computers available, with Control Data Corporation 1604's and IBM 7090's dominating the large machine field. The holder of the unofficial title of the most powerful computer on a college campus has changed many times during the period under review. The current title holder is the Control Data 3600 at Michigan State University. As of the most recent survey, there were 299 computers either installed or on order for colleges and universities (Datamation, 1962). Unfortunately, many of the professional journals have not reflected this change in equipment and have continued to publish far too many programs for IBM 650's and similar obsolete machines. Accompanying the trend to the large-scale "second generation" computers has been a parallel trend to such small, less expensive computers as the IBM 1620 and the Control Data Corporation 160-A. The wide distribution and relatively low operating costs of these small computers have enabled large numbers of persons to obtain firsthand experience in the use of computers. As a result, a market now exists for textbooks written around the design, operation, and programming of a particular computer. The text by Leeson and Dimitry (1962) for the IBM 1620 is an excellent example of this type of book, and similar efforts based upon the popular large-scale computers are certainly in order.

A few years ago, a computer specialist speaking at an educational conference devoted the major portion of his presentation to explaining the characteristics of computers and a minor portion to discussing the relevant research. Since the expanding availability of computers and the prominence of their role in research have resulted in an increased level of computer sophistication on the part of the audience, this introductory phase is no longer necessary. The rise in the level of sophistication of the audience is best exemplified by the increasing number of educational and psychological journals which publish computer-oriented research. Because of the shift in the publication pattern, this writer has attempted to restrict the present chapter to sources easily available to the educational researcher. Therefore, neither articles in the field of engineering nor mimeographed research reports are included. However, the omission does pose some limitations, as many new developments first appear in these sources.

Developments in Computer Usage

Early in the period under review, the educational researcher encountered considerable difficulty in locating adequate generalized programs for the classical statistical analyses, such as analysis of variance, multiple regression, and factor analysis. At the present time, this problem is essentially nonexistent—at least with respect to the popular computers. The excellent BIMED series of statistical programs reported by Dixon (1962) has been adopted at a number of installations. In addition, Leone and others (1961) published an extensive listing of the statistical programs available for many different computers.

Since the need for fundamental statistical library programs has been met for the most part, the current programming effort has tended to be aimed at either improving existing techniques or extending the range of statistical programming. For example, in the former case, Walsh (1961) developed a technique for handling missing data in multiple regression, and Eyman, Dingman, and Meyers (1962) compared several of the techniques available for the rotation problem of factor analysis. In the latter case, Cooley and Lohnes (1962) published a book of formula translation (FORTRAN) programs for a number of techniques in multivariate analysis which are of particular interest to behavioral scientists.

With the basic statistical programming needs met, some effort has been devoted to programs designed to perform analyses peculiar to education and psychology. For example, Baker (1963) developed a generalized item and test analysis program that scores each test paper consisting of multiple choice items and yields the following information: Hoyt's coefficient of reliability based on analysis of variance procedure, frequency distribution of test scores, summary statistics of the sample, and complete item analysis of all item response choices. In a clinical setting, Rome (1962) automated the scoring of answer sheets of the *Minnesota Multiphasic Personality Inventory* (*MMPI*). Kleinmuntz and Alexander (1962) devised a program which performs a diagnosis of *MMPI* profiles based upon the Meehl-Dahlstrom approach.

Noncomputational Uses of Computers

Although just emerging from infancy, the noncomputational uses of computers are assuming a significant role in computer-oriented research. Noncomputational usage is not clearly defined at present; however, what is suggested by such terms as simulation, pattern recognition, information retrieval, and teaching machines would encompass the major aspects of noncomputational applications. The expanse of only a part of the field could be demonstrated by a bibliography of articles on simulation of cognitive processes which itself would readily exceed 500 entries.

The majority of published research in psychological simulation has employed the fifth version of an information-processing language approach known as the IPL-V programming system. Reported in a volume by Newell and others (1961), this procedure has been the outgrowth of the pioneering work of Newell, Shaw, and Simon (1958). This interpretative programming system, which manipulates lists of symbols rather than numbers, represents a considerable departure from ordinary scientific computer programming. This kind of system is, however, ideally suited for simulation.

Simulation in Learning, Problem Solving, and Concept Attainment

The Elementary Perceiving and Memorizing (EPAM) program used by Feigenbaum and Simon (1963) is an IPL-V program designed to perform rote-memory tasks such as memorizing lists of nonsense syllables—a program using the same mechanism as employed by human subjects. Therefore, terms such as *learning* and *memorizing*, customarily employed to describe activities of human subjects, will be taken to represent the humanlike processes achieved by programs for computers used in simulation problems. The EPAM program depends upon a small amount of initial information by which the symbols are paired; during the simulated experiment, this linkage is strengthened until the list is learned to some criterion of performance. In other experiments, the EPAM program has learned to pair symbols representing the sound of an object with the properly spelled name of the object. It should be noted, however, that despite the authors' rather lengthy discussion of EPAM's ability to match "auditory" and "visual" stimuli to the proper object, the computer program involved does treat these lists of symbols as if they were just another list of symbols to be matched. Representation by the lists of both the proper spelling of the object's name and a phonetic spelling is immaterial to the program.

Laughery and Gregg (1962) used IPL-V to simulate the problem-solving behavior of a single human subject who was given as a task the determination of the pattern in which four switches were being set. After considerable study of the "thinking aloud" protocols and the problem-solving behaviors of the subject, an IPL-V program was written to simulate the particular person. Both the human subject and the IPL-V program were then presented with an additional switch-setting problem. The subject and the IPL-V program generated protocols and problem-solving behaviors which exhibited a high degree of similarity. Laughery and Gregg's article contained an explicit account of the painstaking process of developing a simulation program.

From a programming point of view, the simulation of concept attainment is very similar to the type of problem programmed by Laughery and Gregg. Several IPL-V programs have been written for the simulation of concept attainment. Hovland and Hunt's (1960) first program restricted the difficulty of the concept attainment task by providing the subject with a considerable amount of information about the stimuli presented as well as about the concept to be learned. In a second version of the program, the subject had the ability to pay selective attention to certain aspects of the stimulus; in addition, the number of attributes employed in the hypothesis could be varied if additional attributes were needed to attain the concept. The article by Hovland and Hunt, which appears to be one of the earliest presentations of a computer program for concept attainment, would have been greatly enhanced by a program flow chart and by presentation in detail of some actual concept attainment runs. On the basis of their experience, Hovland and Hunt (1960) suggested the inclusion of some additional capabilities in the IPL-V system to simplify programming of concept attainment problems.

An area of concern in studies of concept attainment has been the search strategy (conservative focusing or gambling focusing) employed by the subject. Wickelgren (1962) wrote an IPL-V concept attainment program employing the "conservative focusing" strategy. Rather than assuming that the subject has complete memory of all information, as was the case in the program proposed by Hovland and Hunt (1960), Wickelgren assumed that the subject forgets blocks of data. The article by Wickelgren suffered from lack of detail concerning the actual implementation of the process. Since, in programs such as these, many of the coding decisions are "programmers choice," they may influence the attainment of concepts far more than may the basic strategy employed. For example, if the list of stimuli should be searched in a linear fashion, the speed of attainment could well be a function of position, not strategy. The book by Hunt (1962) provides an excellent overview of both the theoretical and the practical implications of treating concept attainment as an information processing problem. For those readers unfamiliar with literature on concept attainment and its simulation, the book is a helpful starting point.

Simulation of Behavior in Social Groups

The use of the computer to analyze and simulate the behavior of social groups is a new development that should be of great importance to educational research. The analysis of sociometric data is a particularly persistent problem and has been since the inception of the sociogram technique, especially in those situations in which large samples are employed.

Coleman (1961) developed a number of related computer programs for analyzing sociometric data for a number of different purposes. One program employed a scheme for reducing a complex net of sociometric data to the prominent cliques. A second program ascertained the role of reference subjects upon the sociometric choices a subject makes. A third program employed the second to perform a prediction of the role of reference subjects in determining the direction of development of a social system. By means of the third program, one could ascertain whether a student body would become oriented more toward academic subjects than toward athletics over a period of time. He also suggested that stability and instability in three-person groups could be simulated with a computer program.

A social-behavior model has been the basis for an IPL-V simulation program of two-person interaction (Gullahorn and Gullahorn, 1962). The propositions of the model were translated into IPL-V subroutines, which represented the basic information-processing mechanisms of the two subjects. Although both subjects had at their disposal identical mechanisms, the behavior of each subject was unique because of the symbols, the attributes, and the attribute values in the lists possessed by each. Whereas Laughery and Gregg's (1962) program simulated a particular subject and the EPAM program implemented a particular psychological theory, the IPL-V program termed HOMUNCULUS, which was prepared by Gullahorn and Gullahorn (1962), was a blend of these two approaches. Like the EPAM program, HOMUNCULUS rests heavily upon a theoretical basis, and yet it is behavioristic in nature as is Laughery and Gregg's program.

Simulation of Personality

In June 1962, a conference on computer technology and personality theory was held in Princeton at the Educational Testing Service and Princeton University. In addition to serving the purpose of accelerating the dissemination of information regarding computer simulation among personality theorists, the conference was planned to attain some evaluation of the potentialities that computer simulation holds for the field of personality. The contributions of 19 of the participants were published in a book edited by Tomkins and Messick (1963). Needless to say, the volume has great heuristic value. The inclusion of a comprehensive bibliography of approximately 125 references is also noteworthy.

Distinction Between Simulation and Synthesis

Feldman (1962) distinguished between simulation and synthesis. In simulation, the computer program uses the same mechanisms as the original species to perform the task. In synthesis, for a given set of inputs the computer yields outputs identical to those of the species, although it does not use the same mechanisms. The point at which synthesis ceases and simulation begins should generate considerable discussion during the next few years.

Pattern Recognition and Generation

The scholarly discussion by Uhr (1963) of the major developments in pattern recognition devices and programs and of their implications for psychological theory cannot be summarized without injustice to the author. Uhr's paper is required reading for those interested in the field of pattern recognition; its 95 references are an excellent bibliography. In a related article, Selfridge and Neisser (1960) provided a well-illustrated explanation of how a pattern recognition program operates.

In contrast to the use of the computer to recognize patterns, Green (1961) wrote a computer program which (a) generated patterns, (b) displayed them on a cathode ray tube, and (c) photographed the patterns. The photographs were subsequently used as stimuli in perception experiments with human subjects. Computer generation of stimuli provides a means for precise control of the characteristics of stimuli, such as form, angle of rotation, and degree of degeneration required by certain types of experimentation.

Automated Classroom

The proceedings of an October 1961 symposium on the computer-based teaching machine were published in a book edited by Coulson (1962). At that time, the current state of the art of computer-based teaching machines was presented; since then, no significant developments have occurred. The Computer-Based Laboratory for Automatic School Systems, known as CLASS, was constructed by System Development Corporation. An automated classroom was developed that employed a medium-scale computer to control the simultaneous presentation of programed material to as many as 20 pupils and a teacher. The student response stations contained not only cathode ray tubes for presenting material, which were controlled by the computer programs, but also buttons by which the subject communicated his item response choice to the computer. A system somewhat more flexible in design than the CLASS system was developed by the Coordinated Science Laboratory of the University of Illinois and called PLATO II. However, it has not been developed to so high an operational level as the CLASS system.

Although a number of other smaller systems were also presented, the two described exemplify the highest level of development in this field. The missing ingredient in all of the computer-based teaching machine developments has been a concurrent educational research program. It has not been demonstrated that the computer-based teaching machine is more effective as a learning experience than are traditional elementary school workbooks. Additional research should seek those situations in which the unique capabilities of the computer-based teaching machine can make maximum contribution.

Simulation in Educational Administration

Although computer simulation of systems such as fiscal needs, personnel utilization, and determination of plant location is commonplace in industry, work pertaining to analogous simulation of educational systems has not been published. There are many aspects of educational administration which appropriately could be studied via computer simulation: for example, the simulations of automobile traffic performed by the National Bureau of Standards (Communications of the ACM, 1962) might provide the basis for the simulation of pupil flow to be used in the design of school buildings. Hopefully, the next few years will produce some results in this interesting area.

Information Storage and Retrieval

Information storage and retrieval, the so-called library problem, has been an area of continuing effort for a number of years. In a survey of the existing information-retrieval systems, Grems (1962) reported that 3 programming languages have been developed especially for use in information retrieval, that 10 information storage and retrieval systems have been brought to the operational stage, and that 7 systems are considered to be in the research or experimental stages.

One of the most interesting of the many information-retrieval systems is called Baseball. Devised by Green and others (1961), it allows one to present natural text questions about the history of baseball to the system: for example, How many times did the Red Sox play in July 1920? In the Baseball system, the history of baseball is stored within the computer as lists, and these are manipulated by means of an IPL-V program.

Because of the limitations of present computers, documents or abstracts must be in compact form in information-processing systems. The reduced form usually consists of key words or indexes that purport to represent the content of the document. Maron and Kuhns (1960) developed a "probabilistic indexing" scheme in which each key word is assigned a relevance number. This number indicates the degree to which the key word or index relates to the content of the document. They found that the use of such a relevance number improved the retrieval of documents appropriate to a request.

Through means of a factor analysis of key words, Borko (1962a) developed an automatic classification of a corpus of psychological documents. The documents were assigned to categories on the basis of their factor loadings, and the resulting classifications were compared with the classifications of the documents affected through use of the system developed by the American Psychological Association. Only a moderate degree of agreement was found between the two sets of classifications of the documents.

The research effort concerning the whole area of the library problem has been marked by considerable discussion, tentative proposals, and use of heuristic schemes, rather than by generally useful operational systems. A theory of information retrieval is needed to provide a conceptual framework from which adequate information storage and retrieval systems can be developed—hopefully, in time to cope with the publication explosion.

Theory Development

As the noncomputational uses of computers have developed, it has become increasingly apparent that the computer is going to have considerable impact upon the development of psychological and social theories. Because of the small-step processes by which computers function, the theory must be specified precisely—not a trivial task. The nebulous theories prevalent in the educational world cannot survive the cold realities of programming for a digital computer. As a case in point, Uhr (1963) stated that major difficulties have been encountered in recent efforts to put the theories of Hebb and Hull into a precise form that would permit them to be tested. Uhr also indicated that computer programs are on the verge of becoming extremely important theoretical tools and that, unless psychologists take an active role in program construction, they will be excluded from an important development in their own field. This caution can be directed with even greater emphasis to education, for which the majority of theoretical developments are the products of persons outside the field.

As computer programs become increasingly prevalent vehicles for expressing psychological theories, an increased burden of documentation will be placed upon those publishing accounts of computer programs. At the present time, the authors of articles involving IPL-V programs, with a few exceptions, have been negligent both in indicating the assumptions made by the programmer to implement the problem and in providing a detailed account of the course of the simulated experiment. Without such documentation, effective evaluation of the theory involved is not possible.

Developments in Programming Languages

There has been a line of continuing development in programming languages; it has included absolute coding, regional assignment coding, symbolic assembly programs, and the so-called machine independent languages.

Bromberg (1963) surveyed the existing programming languages and found examples of all levels in use at the present time, although the absolute and regional assignment systems, which have been supplanted by sophisticated symbolic assembly routines, are presently considered obsolete. The latter, in turn, are being supplanted by the "higher-level" languages involving formula translation, such as FORTRAN. After a rather slow beginning, the FORTRAN language, described by McCracken (1961), has become the standard scientific programming language of the large-scale computers. It is also available for a number of small machines. Because of the limitations of FORTRAN, many programmers prefer such systems as IBM's Fortran Assembly Program (FAP) or Control Data's Fortran Symbolic, which allows one to combine in a single program the simplicity of the FORTRAN language and the programming flexibility and power of the symbolic assembly form of machine language.

A considerable amount of material has been published concerning the development of an algorithmic computer language (ALGOL), which could serve as a programming language and as a standard publication language. After a number of years of discussion, this language, ALGOL 60, has not gained wide usage because of unresolved difficulties in its machine implementation (Naur, 1963). However, this effort has a number of offshoots. The language termed MAD is an example (Galler, 1962).

For the next few years, at least, FORTRAN probably will continue to be the closest approximation to a universal programming language. The proliferation of manuals on the fundamentals of FORTRAN programming has enabled one to learn easily the bare facts of FORTRAN but not the means for integrating them into a computer program. The book by Cooley and Lohnes (1962) is unique in that the authors presume that the reader knows FORTRAN and illustrate in detail how to use these fundamentals to develop programs for some rather complex statistical analyses.

Need for Improvement in the Selection and Training of Computer Programmers

The increasing reliance upon computers in all areas of research is creating a universal demand for persons skilled in the use of these machines. If sufficient numbers of programmers are to be available in the future, the skills, traits, and attitudes of an effective programmer must be ascertained. Peres and Arnold (1963) made an initial attempt to identify programmer behavior. A high degree of tolerance for frustration appears to be a major characteristic of skilled programmers. Further inquiry into the identification of programming talent and into development of instruments to predict success in programming would certainly be an appropriate and significant task for educational research.

Several university-level courses, such as those proposed by Baker (1962) and Borko (1962b), have been developed to provide present and future behavioral scientists with a working knowledge of existing library programs, of programming fundamentals, and of the role of computers in research. The book edited by Borko (1962b) on computer applications in the behavioral sciences could serve as the focal point for such a course. Feigenbaum (1962) proposed a course to develop the understanding and skills necessary to use IPL-V in programming simulation problems. Such courses as the above represent a sharp break from the traditional mathematically oriented programming course. They should do much to implement an intelligent integration of computers into educational research.

Another attempt to increase the supply of computer programmers has been to teach a wide range of programming topics to bright high school students (Heller, 1962). Although such courses now exist only in a few high schools, an expansion of this program could considerably increase the supply of skilled programmers.

Summary

The period covered by this REVIEW has been marked by rapid developments on a number of fronts in computer technology and applications. Many of these developments may be expected to be reflected in educational research. The increase in the absolute numbers and in the power of the computers available to educational researchers has far exceeded fondest expectations of a few years ago. The widespread availability of generalized statistical programs has made routine practice of such computational procedures as analysis of variance, factor analysis, and item analysis. The time and effort previously expended upon the mechanics of calculation or upon special case programming now can be devoted to design and execution of research.

The current embryonic interest in educational data processing, as indicated by Yasaki (1963), parallels the growth of business data processing some years ago. Hopefully, educators will profit from the considerable experience that industrial managers have so painfully acquired. Educational data processing, once developed, could influence constructively both information processing in educational research and the instructional process itself. College and university educators are in a particularly advantageous position to help meet the critical demand for programmers—through their development of instruments to predict success in programming, through their design of and participation in experiments aimed at improving methods for teaching programming, and through their influence in requiring courses to be a part of the training of social scientists.

There are many areas of computer applications that will profoundly affect the methodology of educational research as well as educational-psychological theory. The simulation of cognitive behavior, the study of small-group interaction, and the analysis of the nature of social systems—all these offer educators a fertile field within which unique contributions could be made. Unfortunately, since educators have been remiss, persons without a primary commitment to education have been performing the pioneering research. Present trends in theory construction seem to indicate that many future discussions of psychological theory will be centered around the details of a computer program and around the implications of these details relative to the performance of the program. To keep abreast of this development, educational researchers will have to become increasingly well informed about a number of different programming systems.

Bibliography

BAKER, FRANK B. "A Computer Course for the Behavioral Scientist." *Educational and Psychological Measurement* 22: 617-21; Autumn 1962.

BAKER, FRANK B. "Generalized Item and Test Analysis Program—A Program for the Control Data 1604 Computer." *Educational and Psychological Measurement* 23: 187-90; Spring 1963.

BORKO, HAROLD. "The Construction of an Empirically Based Mathematically Derived Classification System." *Proceedings: 1962 Spring Joint Computer Conference.* Palo Alto, Calif.: National Press (850 Hansen Way), 1962. pp. 279-89. (a)

BORKO, HAROLD, editor. *Computer Applications in the Behavioral Sciences.* Englewood Cliffs, N.J.: Prentice-Hall, 1962. 633 pp. (b)

BROMBERG, HOWARD. "Survey of Programming Languages and Processors." *Communications of the ACM* 6: 93-99; March 1963.

COLEMAN, JAMES S. "Analysis of Social Structures and Simulation of Social Processes with Electronic Computers." *Educational and Psychological Measurement* 21: 203-18; Spring 1961.

COMMUNICATIONS OF THE ACM. "Computer Simulation of City Traffic." *Communications of the ACM* 5: 224-26; April 1962.

COOLEY, WILLIAM W., and LOHNES, PAUL R. *Multivariate Procedures for the Behavioral Sciences.* New York: John Wiley & Sons, 1962. 211 pp.

COULSON, JOHN E., editor. *Programmed Learning and Computer-Based Instruction.* New York: John Wiley & Sons, 1962. 291 pp.

DATAMATION. "Computing in the University." *Datamation* 8: 27-30; May 1962.

DIXON, W. J. "BIMED: A Series of Statistical Programs for the IBM 709, Part 1." *Behavioral Science* 7: 264-67; April 1962.

EYMAN, R. K.; DINGMAN, H. F.; and MEYERS, C. E. "Comparison of Some Computer Techniques for Factor Analytic Rotation." *Educational and Psychological Measurement* 22: 201-14; Spring 1962.

FEIGENBAUM, EDWARD A. "An Experimental Course in Simulation of Cognitive Processes." *Behavioral Science* 7: 244-45; April 1962.

FEIGENBAUM, EDWARD A., and SIMON, HERBERT A. "Performance of a Reading Task by an Elementary Perceiving and Memorizing Program." *Behavioral Science* 8: 72-76; January 1963.

FELDMAN, JULIAN. "Computer Simulation of Cognitive Processes." *Computer Applications in the Behavioral Sciences.* (Edited by Harold Borko.) Englewood Cliffs, N.J.: Prentice-Hall, 1962. Chapter 15, pp. 336-59.

GALLER, BERNARD A. *The Language of Computers.* New York: McGraw-Hill Book Co., 1962. 244 pp.

GREEN, BERT F., JR. "Using Computers To Study Human Perception." *Educational and Psychological Measurement* 21: 227-33; Spring 1961.

GREEN, BERT F., JR., and OTHERS. "Baseball: An Automatic Question-Answer." *Proceedings of the Western Joint Computer Conference.* New York: Institute of Radio Engineers (1 E. 79th St.), 1961. pp. 219-24.

GREMS, MANDALAY. "A Survey of Languages and Systems for Information Retrieval." *Communications of the ACM* 5: 43-46; January 1962.

GULLAHORN, J. E., and GULLAHORN, J. T. *A Computer Model of Elementary Social Behavior.* SP-741. Santa Monica, Calif.: System Development Corporation, 1962. 15 pp.

HELLER, GEORGE C. "A Computer Curricula for the High School." *Datamation* 8: 23-26; May 1962.

HOVLAND, CARL I., and HUNT, EARL B. "Computer Simulation of Concept Attainment." *Behavioral Science* 5: 265-67; July 1960.

HUNT, EARL B. *Concept Learning: An Information Processing Problem.* New York: John Wiley & Sons, 1962. 286 pp.

KLEINMUNTZ, BENJAMIN, and ALEXANDER, L. BARTON. "Computer Program for the Meehl-Dahlstrom MMPI Profile Rules." *Educational and Psychological Measurement* 22: 193-99; Spring 1962.

LAUGHERY, K. R., and GREGG, L. W. "Simulation of Human Problem-Solving Behavior." *Psychometrika* 27: 265-82; September 1962.

LEESON, DANIEL N., and DIMITRY, DONALD L. *Basic Programming Concepts and the IBM 1620 Computer.* New York: Holt, Rinehart and Winston, 1962. 368 pp.

LEONE, F. C., and OTHERS. *Abstracts of Statistical Computer Routines.* Report No. 1056. Cleveland: Statistical Laboratory, Case Institute of Technology, 1961. 225 pp.

MCCRACKEN, DANIEL D. *A Guide to FORTRAN Programming.* New York: John Wiley & Sons, 1961. 88 pp.

MARON, M. E., and KUHNS, J. L. "On Relevance, Probabilistic Indexing and Information Retrieval." *Journal of the Association for Computing Machinery* 7: 216-44; July 1960.

NAUR, PETER, editor. "Revised Report on the Algorithmic Language ALGOL 60." *Communications of the ACM* 6: 1-17; January 1963.

NEWELL, ALLEN; SHAW, J. C.; and SIMON, HERBERT A. "Elements of a Theory of Human Problem Solving." *Psychological Review* 65: 151-66; May 1958.

NEWELL, ALLEN, and OTHERS, editors. *Information Processing Language-V Manual.* Englewood Cliffs, N.J.: Prentice-Hall, 1961. 244 pp.

PERES, SHERWOOD H., and ARNOLD, PHIL H. "Identifying Programmer Behavior." *Datamation* 9: 40; January 1963.

ROME, HOWARD P. "Automation Techniques in Personality Assessment." *Journal of the American Medical Association* 182: 1069-72; December 15, 1962.

SELFRIDGE, OLIVER G., and NEISSER, ULRIC. "Pattern Recognition by Machine." *Scientific American* 203: 60-68; August 1960.

TOMKINS, SILVAN S., and MESSICK, SAMUEL. *Computer Simulation of Personality.* New York: John Wiley & Sons, 1963. 325 pp.

UHR, LEONARD. "'Pattern Recognition' Computers as Models for Form Perception." *Psychological Bulletin* 60: 40-73; January 1963.

WALSH, JOHN E. "Computer-Feasible Method for Handling Incomplete Data in Regression Analysis." *Journal of the Association for Computing Machinery* 8: 201-11; April 1961.

WICKELGREN, WAYNE A. "A Simulation Program for Concept Attainment by Conservative Focusing." *Behavioral Science* 7: 245-47; April 1962.

YASAKI, ED. "Educational Data Processing." *Datamation* 9: 24-27; June 1963.

Additional References

ALLEN, MAX. "A Concept Attainment Program That Simulates a Simultaneous-Scanning Strategy." *Behavioral Science* 7: 247-50; April 1962.

BAKER, FRANK B. "Information Retrieval Based upon Latent Class Analysis." *Journal of the Association for Computing Machinery* 9: 512-21; October 1962.

BORKO, HAROLD, and BERNICK, MYRNA. "Automatic Document Classification." *Journal of the Association for Computing Machinery* 10: 151-62; April 1963.

BOTTENBRUCH, H. "Structure and Use of ALGOL 60." *Journal of the Association for Computing Machinery* 9: 161-221; April 1962.

DOYLE, LAUREN B. "Semantic Road Maps for Literature Searchers." *Journal of the Association for Computing Machinery* 8: 553-78; October 1961.

FEIGENBAUM, E. A. "The Simulation of Verbal Learning Behavior." *Proceedings of the Western Joint Computer Conference.* New York: Institute of Radio Engineers (1 E. 79th St.), 1961. pp. 121-32.

GREEN, BERT F., JR. *Digital Computers in Research.* New York: McGraw-Hill Book Co., 1963. 333 pp.

GURK, HERBERT M., and MINKER, JACK. "The Design and Simulation of an Information Processing System." *Journal of the Association for Computing Machinery* 8: 260-70; April 1961.

GYR, JOHN; THATCHER, JAMES; and ALLEN, GEORGE. "Computer Simulation of a Model of Cognitive Organization." *Behavioral Science* 7: 111-16; January 1962.

HARE, A. PAUL. "Computer Simulation of Interaction in Small Groups." *Behavioral Science* 6: 261-65; July 1961.

HOLDIMAN, THOMAS A. "Management Techniques for Real Time Computer Programming." *Journal of the Association for Computing Machinery* 9: 387-404; July 1962.

SIMMONS, ROBERT F.; KLEIN, SHELDON; and MCCONOLOGUE, KERNEN. "Toward the Synthesis of Human Language Behavior." *Behavioral Science* 7: 402-407; July 1962.

STONE, PHILIP J., and OTHERS. "The General Inquirer: A Computer System for Content Analysis and Retrieval Based on the Sentence as a Unit of Information." *Behavioral Science* 7: 484-98; October 1962.

Index to Volume XXXIII

Page citations are made to single pages; these are often the beginning of a chapter, section, or running discussion dealing with a topic.

Additivity: 496

Administration in special education: buildings and equipment in relation to, 11; salary structure and differentials within framework of, 11; transportation as problem in, 12

Analyses and implications of disciplines: English, 248; mathematics, 248; science, 248; social sciences, 248

Analysis of variance and design of experiments, developments in: mixed model—repeated measurement designs, 491; multifactor designs without repeated measurements, 493; philosophical basis for inference from experimental studies, 490; special topics in design and analysis, 495

Appraisal function in relation to: communicating appraisal information, 193; evaluation of test reports, 193; interpreting data, 190; nontest methods, 190; scoring and score conversions, 189; test taking and its preparation, 188

Assessment of teacher behavior and instruction: criterion problem, 422; future research, 433; general references concerning, 416; information system model, research literature viewed within framework of, 420; interpersonal relations in classroom, 424; limitations in findings, 431; methodology, 422; observation, 423; overview of trends in research, 415; scheme for classifying teacher behaviors, 423; teacher effectiveness, 431; theory of instruction and teacher behavior, 419

Assumptions in two-group designs: 497

Authorship, inference concerning: 479

Automated instruction and teaching machines: 271

Autonomy of teachers: 402

Background in special education: comparative aspects of, 7; legislation, 6; philosophy and theory, 5; research theory, 6; services, need for and extent of, 6

Bayesian statistical theory: contributions to statistical inference, 462; in decision theory, 557; in psychological literature, 477

Books in statistical methodology, citation and description of: 474

Campus laboratory school: 377

Career conferences: 210

Career development: general, 197; in women, 200

Certification: levels of responsibilities, 344; state-by-state variations, 343; teacher examinations, 345; trends, 346; variations in, 344

Chart or charter, recent developments in educational discourse: 233

Chemistry: 282

College instruction: research in, 431

Computers in educational research: automated classroom in relation to, 571; developments in computer usage, 567; developments in programming languages, 573; information storage and retrieval, 572; noncomputational uses, 568; pattern recognition and generation, 571; selection and training of computer programmers, need for improvement in, 574; simulation in educational administration, 572; simulation in learning, problem solving, and concept attainment, 568; simulation of behavior in social groups, 569; simulation of personality, 570; theory development, use of computers in, 573

Computers in regression analysis: 505

Concepts and purposes underlying guidance and personnel practice: in colleges and universities, 147; in schools, 147

Conditions of employment and service in elementary and secondary schools: economic status and economic conditions of employment, 386; employment, assignment, and turnover, 385; evaluation, retention, and dismissal, 385; needed research in, 388; staff organization and utilization, 383; teacher

load, 384; teacher morale and job satisfaction, 387; teacher status and role expectations, 382
Contingency tables: 484
Contributions to statistical inference: Bayesian statistical theory, 462; Fisher, 458; Neyman and Pearson, 458; Wald, 460
Contributions to theory: 445
Correlation analysis: measures of association, 506; measures of correlation, 507
Counseling: avowed goals, 179; diagnosis, 183; methods, 183; obtained results, 180; theory construction, 179
Counseling procedures, group: 205
Covariance analysis: 497
Crippling conditions and special health problems, children with: adjustment of, 102; characteristics of, 102; curriculum and methods for, 105; educational provisions for, 104; identification and classification of, 99; incidence of, 101; intelligence of, 103; medical screening of, 99; needed research in study of, 105; psychological screening of, 100; special classes and enrollments for, 104
Criterion problem: in assessment of teaching, 422
Criticisms of educational research: 233
Curriculum and communications media: merchandising and research, 302; technology of teaching, 301
Curriculum components and organization: curriculum organization, 288; English, 284; foreign languages, 287; mathematics, 278; social studies, 287
Curriculum development, dynamics of change in: contributions from related fields, 315; diffusion of innovations, 306; field demonstration, 308; need for further research, 316; redefinition of change, 310; roles, role perception, and personality structure in change, 311; search for models, 304
Curriculum organization: team teaching, 288; track organization, 289
Curriculum planning and development: new and promising areas of curriculum research, 332; new and promising research procedures, 333; studies of critical significance in, 330; theory of, 334
Curriculum research: issues in, 326; problem of design, 324; problem of relational contexts, 322
Data analysis and statistical hypothesis testing: 469
Deaf and hard of hearing: communication, techniques of, 51; hearing aids, 49; hearing tests, 49; incidence and causes of deafness, 48; in other countries, 57; integration into hearing groups, 55; intelligence tests, 51; language development, 53; machines to aid, 53; miscellaneous studies concerning, 56; multiple handicaps in relation to, 56; vocational and social adjustment of, 54
Decision theory: conclusion theory in relation to, 467; discussion of, 556
Decision theory and utility management: basic considerations, 557; central concepts, questions, and difficulties, 558; general references, 557; mathematical expectation, 559; probability, 558
Description of teacher behavior based on: direct observation, 428; inferences, 428
Design: in experiments, 490; in factor analytic studies, 510; in sample surveys, 523
Development of theories of statistical inference and significance tests: 457
Disciplines: analyses and implications of, 247; curriculum source recognized, 245
Dismissal of teachers: 385
Economic status and economic conditions of employment for teachers: 386
Educational discourse, recent developments in: 233
Educational research, criticisms of: 233
Elementary and secondary school programs: issues relating to coordination of pupil personnel services, 159; organization and administration of guidance services, 152; roles and functions of pupil personnel workers, 156
Elementary school teacher education: 370
Emotionally and socially handicapped: dynamic factors in therapeutic education for, 116; identification of, 110; identification of dynamic factors in maladjustment of, 111; incidence and screening of, 110; psychoeducational procedures with, 115; role of school in relation to, 115; school consulta-

tion and liaison with regard to, 117; special classes and related educational provisions for, 115; therapies affecting learning of, 118; training of teachers for, 114

Employment, assignment, and turnover: 385

English: review of studies, 285; theory of language instruction, 284

Evaluation: in elementary and secondary schools, 385; in higher education, 392; of research reviewed in issue, 442

Factor analysis, decisions within single study employing: communality and rank, 515; rotational options, 516

Factor analytic methodology: decisions within a single study, 514; design of studies, 510; selected applications of methods, 518

Factor analytic studies, design of: hypotheses and interpretations, 510; other techniques and models, 513; *P*-, *Q*-, and *R*-techniques, 512

Factor studies: contributions to identification of characteristics in training and selection of teachers, 364

Forces influencing curriculum decisions: organized knowledge, 245; psychological, 268; societal, 254

Foreign languages: 287

Foundations in guidance, counseling, and personnel services: concepts and purposes of, 146; philosophical trends in, 146; society and culture in relation to, 143

Future of research in teacher personnel: 446

Gifted: ability grouping in relation to, 91; acceleration for, 93; achievement, underachievement, and achievement motivation in, 88; characteristics and identification of, 84; community and parental attitudes toward special education for, 90; creativity in relation to, 87; curriculum developments and adaptations for, 89; curriculum studies for, 93; economics of talent, 84; general developments concerning, 94; intellectual characteristics of, 84; major books concerning, 94; needed research concerning, 94; personal and social characteristics of, 85

Group counseling: 205

Group dynamics: applications of, 208

Group interaction: 271

Group procedures: applications of group dynamics, 208; counseling with parent groups, 208; group counseling, 205; occupations units, career conferences, and vocational guidance in regular course work, 210; orientation and special guidance courses, 209

High standards needed in selecting publications for the REVIEW: 443

Higher education programs: in extended education, 166; in four-year colleges, 163; in junior colleges, 166; in relation to issues underlying research needs, 168

Hypothesis testing and statistical inference: 455

Hypothesis testing in relation to statistical methodology: development of theories of statistical inference and significance tests, 457; hypothesis testing and statistical inference, 455; problems in statistical hypothesis testing, 466

Identification and classification of children who are: emotionally and socially handicapped, 110; gifted, 84; visually handicapped, 38

Incidence of children with: crippling conditions and special health problems, 101; deafness and hearing loss, 48; emotional and social handicaps, 110; special handicaps, 99; speech handicaps, 20

Information system model of teaching: external inputs—conditions external to the teacher, 426; internal inputs contributing to instructor behavior, 424; products of instructor behaviors—pupil behaviors, 429; teacher behaviors—outputs of teacher information processing, 427; teacher information processing behavior, 426

Information theory: 535

In-service education of teachers: 373

Instructor, or teacher, behavior: teacher-student variables in, 424

Interpreting data in the appraisal function: counselor vs. statistical prediction, 191; differential predictability, 192; expectancy studies, 190; use of *Strong Vocational Interest Blank*, 191

Issues relating to coordination of pupil personnel services: 159

Latin square designs: 495
Legal status of teachers: certification, tenure, and autonomy, pertaining to, 402; civil rights, regarding, 403; compensation, residual rights, and royalties in relation to television, radio, and programed instruction, 404; retirement, with reference to, 405; teacher-board relations, concerning, 404; teacher-pupil relations, concerning, 404
Likelihood principle and statistical inference: 464
Load for teachers: 384

Mathematics: Educational Testing Service study, 279; Minnesota National Laboratory studies, 279; SMSG research, 280
Mathematics and psychology: 540
Mentally retarded: basic educational problems of, 66; curricular adaptations for, 72; educational provisions for, 70; family and community relationships with respect to, 74; follow-up studies of postschool adjustment of, 73; identification and classification of, 62; issues regarding, 76; learning by, 66; motivation of, 68; motor performance of, 69; programed instruction for, 126; research needed with respect to, 76; screening procedures and population trends for, 64; terminology concerning, 62; test developments for, 63; trends regarding, 76
Methodological problems in program evaluation: criterion problem, 218; in relation to experimental design, 220
Methodology in assessment of teacher behavior and instruction: criterion problem, 422; interpersonal relations in the classroom, 424; observation processes, 423; scheme for classifying teacher behaviors, 423
Mixed model—repeated measurement designs: complex mixed designs involving repeated measurements, 493; one fixed factor—repeated measures on random factor, 491; two fixed factors—repeated measures on random factor under one fixed factor, 492
Models: in general, 240; search for, 304
Models in statistical learning, empirical tests of: conditional dependencies, 548; matching rule, 547; variety of statistics, applied, 548
Models in statistical learning, important extensions: concept formation in children, 550; reaction time, 551; response magnitude, 550; social psychology, 549
Models in statistical learning, types of: incremental, 546; pattern, 544; pattern, generalized, 546; strategy, 546
Morale, opinions, and attitudes of teachers: 409
Multidimensional scaling: applications, 538; individual differences, 538
Multifactor designs without repeated measurements: Latin square designs, 495; randomized block designs, 493; response surface and related designs, 494
Multiple comparisons: contributions to, 485; separation of populations in relation to, 495
Multiple handicaps in relation to: deaf and hard of hearing, 56; visually handicapped, 43
Multivariate analysis of variance: 497

Need and preparation of teachers in higher education: 391
New and promising areas of research: in curriculum, 332; in procedures, 333; in teacher personnel, 445
New and promising procedures of research in teacher personnel: 446
New media in teacher education: 374
Null hypothesis, rejection or acceptance of: 467

Occupational classification: 200
Occupational information, sources of: 201
Occupational psychology: 199
Occupations units, career conferences, and vocational guidance in regular course work: 210
Organization and administration of guidance services in relation to: special services, 154; teacher vs. specialist participation, 155; timing in introducing guidance services, 155; types of counselor assignments, 153; use of data-processing equipment, 154; use of ratios, 153
Organization of special education: at federal level, 8; at local level, 9; at state level, 8; extension of programs and services in relation to, 10; segregated and integrated programs of instruction in relation to, 9

Orientation and special guidance courses: 209

Parametric and nonparametric statistics, selected contributions to: books, in statistical methodology, 474; contingency tables, 484; general contributions to the theory of statistical inference, 475; inference concerning authorship, 479; multiple comparisons, 485; significance tests, 480; theory of estimation, 477; theory of statistical inference, 475
Parent groups, counseling with: 208
Perception-diagnosis theory of teaching: 419
Personality characteristics: in teacher selection, 361
Personality structure in relation to dynamics of change in curriculum: 311
Philosophical basis for inference from experimental studies: 490
Philosophical trends in guidance, counseling, and personnel services: 146
Physics: 281
Power and efficiency: 482
Preparation and professionalization of specialists: 175
Preparation of counselors: 172
Pre-service and in-service education for teachers: campus laboratory school, role of, 377; elementary teachers, concerning, 370; in-service programs, 373; new media, with reference to, 374; needed research in, 378; professional laboratory experiences, concerning, 375; secondary teachers, concerning, 371
Pre-service programs: 370
Problems of design in curriculum research: 324
Problems of robustness, power, and efficiency, miscellaneous contributions to: 482
Professionalization of counselors in educational settings: 174
Program evaluation: control group studies, 216; criterion problem in, 218; follow-up studies, 215; general discussions, 214; methodological problems in, 218; problems of experimental design in, 220
Programed instruction for the mentally retarded: arithmetic studies, 133; automated curriculum study, 135; comparison of classroom and programed learning techniques, 130; confirmation procedures, 131; inductive reasoning study, 135; intelligence and step-size study, 134; language studies in, 126; picture program in teaching sight vocabulary, 127; sentence program using verbal prompts, 129; spelling, 132; types of projects in, 126; use of retarded tutors, 128; word learning, 130
Psychological forces influencing curriculum decisions: learner, 268; learning outcomes, 272; learning process, 269; research design and psychological influences, 273; school learning situation, 270
Psychophysics: 534
Published works: of critical significance in area of teacher personnel, 443
Pupil behaviors: products of instructor behaviors, 429; pupil achievement and learning, 430

Randomized block design: 493
Recruitment: age at making decision to be teacher, 356; factors in decision making, 356
Recruitment and selection of prospective elementary and secondary school teachers: factor studies, 364; recruitment, aspects of, 355; selection, factors pertaining to, 358
Regression analysis: accretion and deletion procedures, 503; *B*-nonlinear models, 502; computer analysis, 505; miscellaneous topics, 505; regression models, 501; *x*-nonlinear models, 502
Regression and correlation: correlation analysis, 506; regression analysis, 501
Relational contexts in curriculum research: 322
Research designs and psychological influences: 273
Research needed concerning: conditions of employment and service in elementary and secondary schools, 388; pre-service and in-service teacher education, 378; teacher certification, supply, and demand, 351; understanding teacher behavior, 433
Research procedures, new and promising: 333
Research theory, nature of: analysis of problems, 235; natural history stage, 236; stages of inquiry, 234

Response surface and related designs: 494

Retention of teachers: 385

Robustness: problems of robustness, power, and efficiency, miscellaneous contributions to, 482; *t*-test with the Wilcoxon-Mann-Whitney *U*-test, comparison of, 480

Roles and functions of pupil personnel workers in relation to: health services, 156; nomenclature of guidance services, 157; psychological services, 157; rôle and function of the counselor, 158; self-perceptions of counselors, 157

Roles and role perception in relation to dynamics of change in curriculum: 311

Salaries and other benefits: in elementary and secondary education, 386; in higher education, 394; for teachers in general, 398; for women teachers, 399; in relation to additional employment, 400

Sample surveys in education, design of: development of statistical theory, 524; Furno's study, 528; general references, 523; limitations of educational applications, 528; Project Talent, 527; sampling as means of identifying information in census, 527; teacher surveys, 527; use of available lists in sampling procedures, 526

Scaling and measurement theory: computer models, 539; history of psychometrics, 534; information theory, 535; mathematics and psychology, 540; multidimensional scaling, 536; psychophysics, 534; reviews and integrative summaries, 539; signal detection, 535

Scheme for classifying teacher behaviors: 423

School learning situation: automated instruction and teaching machines, 271; group interaction, 272; task and method variables, 271; teacher, 270

Science: chemistry, 282; physics, 281

Secondary school teacher education: 371

Selection, factors pertaining to: academic competence, 359; career commitment, 361; criteria, 358; personality characteristics, 361

Selection of personnel specialists: 171

Selection, preparation, and professionalization of specialists: 171

Signal detection: 535

Significance tests: miscellaneous contributions to problems of robustness, power, and efficiency, 482; new tests, 483; parametric versus nonparametric tests, 480; robustness of *t*-test with Wilcoxon-Mann-Whitney *U*-test, 480; statistical assumptions in relation to, 468; statistical inference, use of, for, 465; statistical significance and substantive significance, 467

Social problems in relation to public schools: delinquency, 260; dropouts, 259; economic influences, 262; political problems, 261

Social realities in relation to school situation: family life, 258; peer culture, 257; racial and ethnic differences, 254; religion, 255; socioeconomic class, 256

Social status of teachers in relation to: educational level, 406; images and stereotypes of teachers and teaching, 407; occupation of father, 407

Social studies: 287

Society and culture in guidance, counseling, and personnel services in relation to: college students, 143; culturally deprived, 145; school-age youth, 144; values, 144

Special classes and related educational provisions for the emotionally and socially handicapped: dynamic factors in therapeutic education, 116; school consultation and liaison, 117; therapies affecting learning, 118

Special education: administration, 11; background, 5; needed research in, 15; organization, 8; supervision, 12

Special topics in design and analysis of experiments: additivity, 496; assumptions in two-group designs, 497; covariance analysis, 497; multiple comparisons and separation of populations, 495; multivariate analysis of variance, 497

Speech handicapped: aphasia in, 26; articulatory disorders in, 22; cerebral palsy in, 27; cleft palate in relation to, 27; incidence of, 20; language development in, 21; laryngectomy in, 27; miscellaneous contributions to study of, 30; research on, trends in, 31; speech problems related to other disorders, 29; speech science, 31; stuttering in, 24; testing and research

techniques for, 20; voice disorders in, 23
Staff organization and utilization: 383
Statistical hypothesis testing, problems in: data analysis and statistical hypothesis testing, 469; decision theory and conclusion theory, 467; null hypothesis, rejection or acceptance of, 467; significance tests and statistical assumptions, 468; statistical significance and substantive significance, 467; statistics and experimentation, 469; statistics and psychological theory, 470
Statistical inference: Bayesian statistical theory, contributions of, 462; Fisher, contributions of, 458; hypothesis testing in relation to, 455; likelihood principle in relation to, 464; Neyman and Pearson, contributions of, 458; significance tests, use of, 465; theories of, in relation to significance tests, 457; theory of, general contributions to, 475; Wald, contributions of, 460
Statistical learning models: alterations of older models, 544; empirical tests, 547; extensions, 549; other models, 551; prospectus, 552
Statistics: experimentation in relation to, 469; psychological theory in relation to, 470
Structures and groupings of disciplines: 249
Studies of critical significance in curriculum planning and development: 330
Summary of teacher-student variables: as considered in the literature, 424
Supervision in special education: community and parent relationships in relation to, 14; evaluation of instruction as problem in, 12; guidance, placement, and follow-up, 13
Supply and demand: economic factors, 347; needed research in, 351; remedies for the shortage, 350; subject and grade level factors, 348; other factors, 349
Surveys: *see* sample surveys in education

Task and method variables: 271
Teacher: 270
Teacher and teaching effectiveness: characteristics, 296; personality, 295
Teacher as problem solver: model for, 421
Teacher attitudes and opinions: 409
Teacher-board relations: legal aspects of, 404
Teacher certification, supply, and demand: 343
Teacher education: 369
Teacher effectiveness: 431
Teacher examinations: 345
Teacher load: 384
Teacher morale: 387
Teacher personnel research in higher education: need and preparation of teachers, 391; new data sources concerning, 396; ratings and evaluation in, 392; salaries and other benefits, pertaining to, 394
Teacher-pupil relations: legal aspects of, 404
Teacher status and role expectations: 382
Teacher-student variables in relation to systems approach: 424
Teaching: analysis of teaching act, 293; as system of social action, 419; teacher and teaching effectiveness, 295; technology of, 301
Team teaching: 288
Technology of teaching: 301
Tenure of teachers: 402
Test taking and its preparation in relation to: instructions, 188; perception, 189; practice, 188; screening, 189
Theory for education: defined, 239; in educational research, 238; in relation to curriculum development, 238; usages of the term, 237
Theory of curriculum planning and development: 334
Theory of estimation: interval estimation, 478; point estimation, 477
Theory of instruction and teacher behavior: contributions in general, 445; information system model, 420; model viewing teacher as problem solver, 421; perception-diagnosis theory of teaching, 419; teaching as system of social action, 419
Theory of statistical inference, general contributions: AERA-sponsored publications, 476; mathematical statisticians, 477; psychological literature, 476; three basic contributions, 475
Theory making in education: incomplete analogy, 242; models, 240
Track organization: 289

Visually handicapped: general works concerning, 38; identification and classification of, 38; mobility of, 41; multiply handicapped in relation to, 43; professional preparation of workers with, 44; reading and writing by, 40; social psychology regarding, 42; tests and psychological studies concerning, 41

Vocational development process: career development, 197; occupational classification, 200; occupational information, sources of, 201; occupational psychology, 199; women's career development, 200

Vocational guidance in regular course work, 210